



(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention
of the grant of the patent:
02.12.1998 Bulletin 1998/49

(21) Application number: 91810167.6

(22) Date of filing: 13.03.1991

(51) Int Cl.⁶: **C12N 15/31**, C12N 15/62,
C12N 15/74, C12N 15/15,
C12N 1/21, C12P 21/02
//(C12N1/21;; C12R1:19;; 1:07;;
1:46)

(54) **Bacterial vectors**

Bakterielle Vektoren

Vecteurs bactériens

(84) Designated Contracting States:
AT BE CH DE DK ES FR GB GR IT LI LU NL SE

(30) Priority: 22.03.1990 GB 9006400

(43) Date of publication of application:
02.10.1991 Bulletin 1991/40

(83) Declaration under Rule 28(4) EPC (expert
solution)

(73) Proprietor: **Société des Produits Nestlé S.A.**
1800 Vevey (CH)

(72) Inventors:
• Suri, Bruno, Dr.
CH-4051 Basle (CH)
• Schmitz, Albert, Dr.
CH-4056 Basle (CH)

(56) References cited:
EP-A- 228 726 **EP-A- 455 280**

- FEMS MICROBIOLOGY LETTERS vol. 30, 1985, ELSEVIER. AMSTERDAM; NL pages 193 - 196; M. J. GASSON ET AL: 'High copy number plasmid vectors for use in lactic streptococci'
- NETHERLANDS MILK AND DAIRY JOURNAL. vol. 40, no. 2/3, 1986, WAGENINGEN NL pages 141 - 154; W. M. DE VOS: 'Gene cloning in lactic streptococci'
- JOURNAL OF BACTERIOLOGY vol. 154, no. 1, 1983, BALTIMORE pages 1 - 9; M. J. GASSON: 'Plasmid complements of Streptococcus lactis NCDO 712 and other lactic Streptococci after protoplast-induced curing'
- GENE. vol. 95, September 1990, AMSTERDAM NL pages 155 - 160; M. VAN ASSELDONK ET AL: 'Cloning of usp45, a gene encoding a secreted protein from Lactococcus lactis subsp. lactis'

Note: Within nine months from the publication of the mention of the grant of the European patent, any person may give notice to the European Patent Office of opposition to the European patent granted. Notice of opposition shall be filed in a written reasoned statement. It shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

DescriptionField of the invention

5 The invention relates to the field of genetic engineering and provides novel DNA molecules which comprise a DNA sequence coding for the most abundant protein in the supernatant of cultures of Lactococcus lactis, hereinafter named Major Secretion Product (MSP), for its signal peptide and/or the promoter of the MSP gene. The molecules may also comprise multifunctional origin of replication. The novel DNA molecules are used for the production of novel shuttle vectors for cloning of DNA in at least E. coli and Lactococcus spec. or novel vectors for the expression of homologous or heterologous genes and the production of secreted gene products in gram positive bacteria such as Lactococcus spec. and Bacillus spec.

Background of the invention

15 Although in genetic engineering techniques numerous procaryotic vector-host-systems for cloning of heterologous or homologous genes are already known, there is a continuous need for novel systems which may have advantages over the known systems.

Most recombinant work in DNA technology has been carried out with bacteria such as Escherichia coli or Bacillus subtilis. The lactic acid bacteria, however, are of much more industrial interest. For this reason, a number of efforts have been made to develop plasmid vectors for cloning and expression of homologous and heterologous genes in lactic acid bacteria, particularly in Lactobacillus spec. and Lactococcus spec. [see PCT WO 85/03945, Gasson and Anderson (1985), EP-A-0 316 677, Bates et al. (1989), Jos et al. (1985) and review articles of Chassy (1987) and of De Vos (1987)]. Lactococcus spec. was formerly named Streptococcus spec.

Lactococcal strains investigated so far harbour a characteristic plasmid complement consisting of multiple different plasmids. This property can be used to differentiate between various lactococcal strains (Davies et al., 1981). For genetic studies, plasmid free strains have been constructed by repeated curing in the course of plasmid function studies (De Vos, 1987).

For example, the plasmid complement of L. lactis 712, hereinafter also called L. lactis LL712, consists of 5 plasmids having the molecular weights of 1.8 Md, 2.5 Md, 5.2 Md, 9 Md and 33 Md which are named pSH71, pSH72, pSH73, pSH74 and pLP712, respectively (Gasson, 1983).

Based on plasmid pSH71 of L. lactis and on the related L. cremoris plasmid pWV01 (Otto et al., 1982) various cloning vectors have been constructed. The cloning vectors have been produced either by inserting genetic markers such as antibiotic resistance genes into the plasmids or by screening fragments of the plasmids for an origin of replication function, i.e. for the ability to sustain replication of selected DNA fragments. A cloning vector produced according to the latter method is pNZ12 which contains the 1.7 kbp ClaI restriction fragment of pSH71 comprising the origin of replication (Gasson and Anderson, 1985). The origin of replication of pSH71 is also functional in other gram-positive bacteria such as Bacillus spec. and in the gram-negative Escherichia coli.

On the basis of these plasmids cloning vectors useful for the introduction and expression of homologous or heterologous genes in lactic acid bacteria have been developed. The development of the cloning vectors resulted in transformed lactococcal strains with improved properties which are useful in food and feed industry, for example a bacteriophage resistant L. lactis strain (EP-A-0 316 677) or a L. lactis strain which produces bovine prochymosin (PCT WO 85/03945). The development of cloning vectors for the production of homologous or heterologous gene products is not only of interest because of the production of improved lactic acid bacteria cells but also for the production of recombinant proteins. One of the major problems with the production of heterologous proteins in microbial expression systems has been the purification of the product. Purification of intracellular proteins is time-consuming and often results in poor yields. Purification can be considerably facilitated if the product is secreted from the host cell. To avoid the problems of purification of the products expressed in bacteria, vector-host systems for the production of recombinant proteins which are secreted into the supernatant can be useful.

Another advantage of secreted proteins can be that they can have a native and biologically active conformation, because then no refolding process is needed. Refolding is usually necessary if the polypeptide is intracellularly deposited.

Secretion of a protein usually requires a signal peptide at the amino terminus of the primary translation product which directs the protein into the secretory pathway. It is of advantage if the signal peptide is cleaved enzymatically from the protein during the translocation through the cell membrane. This, however, is not always the case.

Object of the invention and solutions

It is an object of the invention to provide novel hybrid vectors which can replicate in gram positive and gram negative

bacteria and/or which allow expression of homologous or heterologous genes and secretion of stable protein products into the supernatant.

In particular, this object has been achieved by the present novel hybrid vectors comprising a novel DNA insert which comprises the promoter region, the DNA sequence coding for the signal peptide and the coding region of a hitherto unknown gene encoding a polypeptide which is the most abundant protein in the supernatant of *L. lactis* as judged after TCA precipitation of the supernatant, SDS-polyacrylamide electrophoresis of the precipitated proteins and staining of the gel with Coomassie brilliant blue and which is called herein Major Secretion Product (MSP).

Accordingly, another solution to an object of the invention is to provide novel hybrid vectors comprising the promoter region, the DNA sequence coding for the signal peptide and/or the coding region of the MSP gene or of a related gene and an origin of replication derived from the 2.5 Md or 5.2 Md plasmid of *L. lactis* LL712 or an origin of replication related thereto.

The invention concerns also functional fragments of the novel DNA inserts or the origins of replication per se. They are useful for the production of novel expression vectors for the secretion of homologous or heterologous proteins from lactic acid bacteria.

The invention further provides a method for the preparation of the novel DNA molecules and hybrid vectors, and a method for the production of secreted gene products by means of hosts transformed with a novel hybrid vector of the invention.

The invention provides also the MSP protein in pure form.

Detailed description of the invention

DNA molecules

The invention concerns a hybrid vector comprising

a) the approximately 3.5 kbp EcoRI/Sall *L. lactis* insert of the plasmid pUCRS deposited under the accession number DSM 5803, or a functional fragment thereof, or

b) a DNA sequence which hybridizes with said insert or with a functional fragment thereof, or comprises a promoter region which is naturally operatively linked to such a hybridizing DNA sequence, or

c) a degenerate sequence of a DNA sequence which is covered in a) and which encodes a signal peptide, or

d) a derivative of a DNA molecule covered in a), b) or c),

The vector may also comprise the origin of replication of (I) the 2.5 Md plasmid of *L. lactis* LL712 deposited under the accession number DSM 5804, or of (II) the 5.2 Md plasmid of *L. lactis* LL712, or of (III) a plasmid of the same incompatibility group as the 2.5 Md plasmid or the 5.2 Md plasmid of *L. lactis* LL712.

The approximately 3.5 kbp EcoRI/Sall *L. lactis* insert of the plasmid pUCRS comprises the promoter region of the MSP gene, the DNA sequence encoding the MSP signal peptide and the coding region for MSP. MSP with its signal peptide attached is hereinafter named pre-MSP. The EcoRI/Sall insert comprises the 1920 bp long DNA sequence with the SEQ ID No. 1 depicted in the sequence listing in such orientation that base position 1 is proximal to the EcoRI site.

The coding region for MSP extends from base position 492 up to base position 1793 of the sequence with SEQ ID No. 1.

The DNA sequence which encodes the 27 amino acid long MSP signal peptide, extends from the end of the promoter region defined hereinafter up to the beginning of the DNA sequence coding for the mature MSP, i.e. from base position 411 up to position 491 of the sequence with SEQ ID No. 1. The MSP signal peptide or a functional fragment thereof causes a protein to be excreted from a host cell, e.g. from a lactic acid bacterium, such as *Lactococcus lactis*, or from *Bacillus spec.*, such as *B. thuringiensis*, if the N-terminus of the protein is covalently linked to the C-terminus of the signal peptide. Covalent linkage of the C-terminus of the MSP signal peptide and the N-terminus of a protein which is to be excreted from a host cell can be obtained by the production of a fusion gene in which the entire DNA sequence encoding the MSP signal peptide or part thereof encoding a functional fragment of the MSP signal peptide comprising the C-terminus is linked in proper reading frame to the 5' end of a structural gene encoding the desired protein. Such a fusion gene is, for example, the DNA molecule with SEQ No. 2 which comprises the DNA sequence encoding the MSP signal peptide and the hirudin structural gene.

The amino acid sequences of the MSP signal peptide and of MSP are also given in the sequence with SEQ ID No. 1.

The promoter region of the MSP gene is located upstream of the DNA sequence encoding the MSP signal peptide

and comprises up to about 1500, preferably up to about 100 to 1000 nucleotides. In the approximately 3.5 kbp long EcoRI/Sall L. lactis insert of the plasmid pUCRS, the promoter region extends from the EcoRI cut end of the insert located upstream of the base corresponding to position 1 of the DNA sequence with SEQ ID No. 1 up to base position 411. More particular, the promoter region extends from the first HindIII restriction site located upstream of the base corresponding to position 1 up to base position 411. The promoter region can bind RNA polymerase as well as regulatory proteins and can control the expression of a structural gene operatively linked therewith.

The term functional fragment includes those DNA fragments retaining promoter, signal and/or structural functions.

Preferred functional fragments are those containing the MSP promoter region, the DNA sequence encoding the MSP signal peptide or the promoter region and the DNA sequence encoding the MSP signal peptide. Fragments according to the invention are also composed of smaller fragments retaining the MSP promoter activity and/or coding for a peptide with signal peptide activity.

Fragments with promoter function, for example, are those starting with the first base at the EcoRI cut end of the approximately 3.5 kbp EcoRI/Sall L. lactis insert or, in particular, those starting at the HindIII site and extending up to a base corresponding to about position 410 of the sequence with SEQ ID No. 1. Other fragments with promoter function are selected from the group of fragments starting with any one of the bases between the EcoRI and the HindIII site and ending with a base in about position 410 of the sequence with SEQ ID No. 1. Shorter fragments of the promoter region also retain promoter activity.

A DNA fragment which encodes a signal peptide extends, for example, from about base 411 up to base 491 of the sequence with SEQ ID No. 1. It can be extended at the 5' end with a fragment of the promoter region which does not retain promoter activity.

A fragment retaining MSP promoter function and encoding a signal peptide extends from the EcoRI cut end of the insert located upstream of the base corresponding to position 1 of the DNA sequence with SEQ ID No. 1 down to base position 491. Another fragment retaining MSP promoter function and encoding a signal peptide extends from the first HindIII restriction site located upstream of the base corresponding to position 1 of DNA sequence with SEQ ID No. 1 up to about the base corresponding to position 491. Other fragments retaining MSP promoter function and encoding a signal peptide are selected from the group of fragments starting with any one of the bases between the said EcoRI and HindIII sites and ending with a base in about position 491 of the sequence with SEQ ID No. 1.

The fragments may contain linkers which provide for successful linkage to other DNA molecules. Suitable linkers to above fragments have a DNA sequence which fits into the restriction site of the DNA to which the fragment is to be linked. They may also contain a predetermined restriction site.

A DNA molecule which hybridizes with said insert is hybridizing under conventional conditions. A conventional hybridization procedure is described e.g. by Benton and Davis (1977). Such a hybridizing DNA molecule comprises, for example, as hybridizing DNA sequence a variant of the structural gene encoding MSP. Such a hybridizing variant is hereinafter also referred to as a structural gene related to the MSP gene. The protein encoded by such a related gene is referred to as a protein related to MSP. Accordingly, a promoter region which is naturally operatively linked to a DNA sequence hybridizing with said insert is a promoter region of a gene related to the MSP gene.

A structural gene related to the MSP gene is, for example, a naturally occurring variant derived from another bacterium than L. lactis, particularly from another lactic acid bacterium, for example from Lactobacillus spec. or Lactococcus spec., e.g. L. cremoris or L. thermophilus. It is also a naturally occurring variant which is encoded by an isogene on the chromosome of or on a plasmid naturally occurring in L. lactis.

A DNA molecule comprising a gene related to the MSP gene can be isolated according to conventional methods.

DNA molecules of the invention are also such having degenerated DNA sequences. They are degenerated within the meaning of the genetic code in that an unlimited number of nucleotides are replaced by other nucleotides without changing the amino acid sequence for which they code. Molecules having such degenerate DNA sequences may be useful because of their different restriction sites or because of a preferred codon usage in a particular host. Preferred are degenerated sequences of the DNA sequence coding for the MSP signal peptide or of a variant thereof.

The term derivative when used in connection with a DNA molecule covered in a), b) or c) includes fragments, mutants or larger derivatives of such a DNA molecule. The term derivative includes also larger derivatives of fragments or mutants of such a DNA molecule. Preferred are fragments retaining promoter, signal or structural functions. Examples of such fragments are given above.

A mutant of a DNA molecule described in a), b) or c) is, for example, a DNA molecule having a deletion, insertion, inversion, or point mutation which may occur naturally or may be artificially introduced into the DNA molecule in vivo or in vitro according to conventional methods. A mutant can show an altered restriction pattern.

Larger derivatives of a DNA molecule covered in a), b) or c) are, for example, excisable from the L. lactis genome. They can be found in a genomic library of L. lactis LM0230 obtained by fragmentation of the nucleic acids, treatment of the fragments with a suitable restriction enzyme, e.g. Sau3AI, EcoRI, BamHI or HindIII, ligating into a suitable vector, e.g. the lambda phage λ EMBL3 or the plasmid pBR322, cloning, e.g. in E. coli, and excising again, with the same or another suitable restriction enzyme.

Larger derivatives of a DNA molecule covered in a), b) or c) are also recombinant DNA molecules with flanking sequences, for example such comprising linkers which provide suitable restriction sites or which put regulatory sequences, such as a promoter, a DNA sequence encoding a signal peptide or a terminator, and a structural gene into the correct distance or reading frame, or such comprising sequences derived from a vector, e.g. from a phage or plasmid used in the construction of a hybrid vector. The flanking sequences comprised in the larger derivatives may give rise to fusion genes.

An origin of replication derived from the L. lactis LL712 2.5 Md or 5.2 Md plasmid or from a plasmid of the same incompatibility group as the 2.5 Md plasmid or the 5.2 Md plasmid can cause a plasmid to replicate in acid bacteria, for example in lactococci, e.g. in L. lactis, L. lactis diacetylactis or L. cremoris, and in E. coli or also in gram negative bacteria other than E. coli or in gram positive bacteria other than lactic acid bacteria, e.g. in Bacillus spec. such as B. thuringiensis.

An origin of replication which is derived from the L. lactis LL712 2.5 Md or 5.2 Md plasmid or from a plasmid of the same incompatibility group as the 2.5 Md plasmid or the 5.2 Md plasmid is, for example, comprised in the whole respective linearized plasmid obtainable after treatment with a suitable restriction endonuclease. Accordingly, a hybrid vector of the invention is also such comprising the whole DNA sequence of the 2.5 Md or the 5.2 Md plasmid or of a plasmid of the same incompatibility group as the 2.5 Md or 5.2 Md plasmid. Both plasmids can be isolated, for example, from the strain L. lactis LL712 which is deposited with the DSM as DSM 5804.

A DNA sequence which functions as an origin of replication is also a fragment of any of said plasmids retaining origin of replication function. Such a DNA fragment can be obtained, for example, by isolating it after the fragmentation of the respective plasmid or of a fragment thereof, for example, with physical forces, e.g. shearing forces, or with chemical cleavage reactions, or with enzymes cleaving DNA such as nucleases, e.g. the exonucleases Bal31, S1 or exonuclease III, or endonucleases, e.g. restriction endonucleases. In particular, such a DNA fragment is a restriction fragment obtained after treatment of the respective plasmid or fragment thereof with one or with two different restriction endonucleases which are in the case of the 2.5 Md plasmid, for example, NdeI, SphI and/or EcoRV and in the case of the 5.2 Md plasmid, for example, Sau3A, NdeI, HindIII, AclI and/or EcoRI.

The origin of replication of the 2.5 Md plasmid is located, for example, on the approximately 1.8 kbp long NdeI/SphI fragment, the approximately 3.1 kbp long EcoRV/SphI fragment obtainable after partial EcoRV digestion of the approximately 3.5 kbp long SphI fragment, the approximately 3.5 kbp long SphI fragment, the approximately 1.2 kbp long NdeI/EcoRV fragment, the approximately 2.5 kbp long EcoRV fragment, or the approximately 3.5 kbp long EcoRV/SphI fragment obtainable after partial EcoRV digestion of the approximately 3.5 kbp long SphI fragment. The approximately 3.5 kbp long SphI and the other fragments mentioned hereinbefore are also comprised in pSC12. The production of pSC12 is described in Example 1.2. A restriction map is given in Figure 2.

The origin of replication of the 5.2 Md plasmid is located, for example, on the approximately 1.0 kbp long NdeI/HindIII fragment, the approximately 2.2 kbp long NdeI/AclI fragment, the approximately 2.7 kbp long NdeI/EcoRI fragment, the approximately 3.1 kbp long NdeI/Sau3A fragment, the approximately 2.0 kbp long Sau3A/HindIII fragment, the approximately 3.2 kbp long Sau3A/AclI fragment, the approximately 3.7 kbp long Sau3A/EcoRI fragment, or the approximately 4.1 kbp long Sau3A fragment. The fragments with Sau3A cut ends are obtainable after partial Sau3A digestion of the 5.2 Md plasmid, of pSC18 or of suitable restriction fragments of these plasmids. The production of pSC18 is described in Example 1.3. A restriction map is given in Figure 3.

Plasmids of the same incompatibility group cannot stably coexist in the same cell. They carry related origins of replication.

A recombinant DNA molecule does, for example, contain linkers and/or sequences derived from a vector such as from a phage or plasmid, and optionally a marker gene such as a resistance marker gene, e.g. an erythromycin, ampicillin, or tetracyclin resistance marker gene or the like. A recombinant DNA molecule comprising an origin of replication of the present invention is, for example, a vector plasmid for lactic acid bacteria, preferentially a shuttle vector which can replicate at least in lactic acid bacteria and E. coli and wherein a lactococcal derived origin of the invention is the only origin of replication.

A preferred shuttle vector is pSC12, pSC12ΔP, pSC12ΔN, pSC12ΔNP, pSC18, pSC18ΔP, or pSC18ΔN. A preferred shuttle vector for the expression of a foreign gene in lactic acid bacteria is pSC12HIR, pSC12HIR-1 or pSC12HIRTerm. The preferred shuttle vectors are described in the Examples.

A DNA sequence which functions as an origin of replication in at least lactic acid bacteria and E. coli which is intended to be covered above under e) is also an inversion, insertion, deletion or point mutant of a DNA sequence defined hereinbefore which retains the function of an origin of replication. Such mutants have an altered nucleotide sequence in the DNA sequences flanking the real origin of replication, for example point mutations which cause an altered pattern of restriction sites or deletion mutants such as the L. lactis inserts of the plasmids pSC12ΔP, pSC12ΔN, pSC12ΔNP, pSC18ΔNP, or pSC18ΔN (see Table 1). The real origin of replication mentioned hereinbefore is intended to be the smallest DNA fragment which functions as an origin of replication.

The hybrid vectors of the invention are usable for cloning in hosts such as fungi or in particular bacteria. They can

be derived from any vector useful in the art of genetic engineering, such as from phages, cosmids or plasmids. They are, for example, derivatives of phage λ , e.g. NM 989 or EMBL 3, of phage M13, e.g. M13mp18 or M13mp19, of bacterial plasmids, e.g. pBR 322, pUC18 or pUC19 or plasmids of lactic acid bacteria, or of yeast plasmids, e.g. yeast 2 μ plasmid, or also of defective phages or defective plasmids in the presence of a helper phage or a helper plasmid allowing replication of said defective phages or plasmids.

A hybrid vector of the invention comprises a nucleotide sequence of a DNA molecule covered in a), b), c) or d). It comprises in particular the approximately 3.5 kbp EcoRI/Sall L. lactis insert, a fragment thereof coding for the MSP signal peptide and/or the promoter of the MSP gene. Depending on the type of the DNA molecule of the invention inserted in the hybrid vector, the vector may comprise a hybrid expression control sequence. Such a hybrid expression control sequence partially consists of a DNA molecule covered in a), b) or c) and partially of a DNA molecule different from a DNA molecule covered in a), b) or c). For example, the hybrid vector can comprise the MSP promoter in combination with a DNA sequence coding for a non-MSP signal peptide a DNA fragment coding for the MSP signal peptide in combination with another than the MSP promoter. Additionally, a hybrid vector of the invention optionally comprises a transcriptional terminator region, e.g. the trpA transcription terminator of E. coli.

Hybrid vectors comprising a DNA sequence coding for a signal peptide of the present invention, i.e. derived from the gene for pre-MSP or a related gene, are useful for the production of secreted gene products in lactic acid bacteria, particularly in Lactococcus spec., especially in L. lactis or L. cremoris or in other bacteria, for example a Bacillus spec. such as B. thuringiensis. Such a hybrid vector comprises a promoter which functions in the desired bacterium, for example the MSP promoter for the expression in lactic acid bacteria, and optionally a transcription terminator region which functions in the desired bacterium. The use of a terminator can increase the yield of the recombinant gene product if it is operatively linked with the respective homologous or heterologous structural gene. For example, if the trpA terminator of E. coli is operatively linked with a hirudin gene (EP-168 342) which is expressed under the control of the MSP promoter and a DNA molecule encoding the MSP signal peptide in lactic acid bacteria, the production of desulfatohirudin (vgl. EP-142 860) is markedly increased.

A hybrid vector of the invention also comprises an origin of replication which functions in the desired host, e.g. in lactic acid bacteria and/or Bacillus spec. and/or E. coli.

A DNA sequence bearing an origin of replication which functions in at least lactic acid bacteria is, for example, the 1.7 kbp ClaI fragment of the lactococcal plasmid pSH71 (Gasson and Anderson, 1985) or one of the DNA fragments defined hereinbefore which comprise the origin of replication of the 2.5 Md plasmid or of the 5.2 Md plasmid of L. lactis LL712.

A hybrid vector of the present invention may contain a selectable marker. The choice of the marker depends on the host which is to be transformed, selected and cloned. Any marker gene can be used which facilitates the selection of transformants due to the phenotypic expression of the marker. Suitable markers are particularly those expressing antibiotic resistance, e.g. against erythromycin, tetracycline or ampicillin, or in the case of auxotrophic mutants, genes which complement host lesions, e.g. genes of the lactose metabolizing pathway.

Preferred embodiments of the present invention are hybrid vectors comprising a homologous or heterologous structural gene, for example the MSP or especially a heterologous structural gene defined hereinafter, which is operatively linked in a proper reading frame to the DNA sequence encoding the MSP signal peptide and which can be expressed under the control of the MSP promoter or a heterologous promoter.

Structural genes may originate from viruses, procaryotic cells or eucaryotic cells and may be derived from genomic DNA or from cDNA prepared via the mRNA route, or may be synthesized chemically. They may code for a wide variety of useful polypeptides, including glycosylated polypeptides, in particular of higher eukaryotic, especially mammalian, such as animal or especially human origin, such as enzymes which can be used, for example, for the production of nutrients and for performing enzymatic reactions in chemistry, or polypeptides which are useful and valuable for the treatment of human or animal diseases or for the prevention thereof, for example hormones, polypeptides with immunomodulatory, anti-viral and anti-tumor properties, antibodies, viral antigens, vaccines, clotting factors, foodstuffs, and the like.

Examples of such heterologous structural genes are e.g. those coding for hormones such as secretin, thymosin, relaxin, calcitonin, luteinizing hormone, parathyroid hormone, adrenocorticotropin, melanocyte-stimulating hormone, β -lipotropin, urogastrone or insulin, growth factors, such as epidermal growth factor, insulin-like growth factor (IGF), e.g. IGF-I and IGF-II, mast cell growth factor, nerve growth factor, glia derived nerve cell growth factor, or transforming growth factor (TGF), such as TGF β , growth hormones, such as human or bovine growth hormones, interleukin, such as interleukin-1 or -2, human macrophage migration inhibitory factor (MIF), interferons, such as human α -interferon, for example interferon- α A, α B, α D or α F, β -interferon, γ -interferon or a hybrid interferon, for example an α A- α D- or an α B- α D-hybrid interferon, especially the hybrid interferon BDBB, proteinase inhibitors such as α_1 -antitrypsin, SLPI and the like, hepatitis virus antigens, such as hepatitis B virus surface or core antigen or hepatitis A virus antigen, or hepatitis nonA-nonB antigen, plasminogen activators, such as tissue plasminogen activator or urokinase, tumour necrosis factor, somatostatin, renin, β -endorphin, immunoglobulins, such as the light and/or heavy chains of immunoglobulin D, E or

G, or human-mouse hybrid immunoglobulins, immunoglobulin binding factors, such as immunoglobulin E binding factor, e.g. sCD23, calcitonin, human calcitonin-related peptide, blood clotting factors, such as factor IX or VIIIc, erythropoietin, eglin, such as eglin C, hirudin, desulfatohirudin, such as desulfatohirudin variant HV1, HV2, HV3, [Leu¹, Thr²]-HV1, [Lys⁴⁷]-HV2 or PA, human superoxide dismutase, viral thymidin kinase, β -lactamase, glucose isomerase. A preferred gene is this coding for desulfatohirudin, e.g. variant HV1. In the hybrid vectors of the present invention, the present promoter and/or the DNA molecule encoding the MSP signal peptide is operably linked to the polypeptide coding region so as to ensure effective expression of the polypeptide.

Preferred hybrid vectors are pUCRS, M13 mp18RS, M13 mp19H, pVAHIR, pVAHIR-1, pSC12HIR, pSC12HIR-1, and in particular pSC12HIRTerm. The hybrid vectors are described hereinafter in the Examples.

The invention also concerns a DNA molecule which

a) is the approximately 3.5 kbp EcoRI/Sall L. lactis insert of the plasmid pUCRS deposited under the accession number DSM 5803, or a functional fragment thereof, or

b) hybridizes with said insert or with a functional fragment thereof, or comprises a promoter region which is naturally operatively linked to such a hybridizing DNA sequence, or

c) is a degenerate sequence of a DNA sequence which is covered in a) and which encodes a signal peptide, or

d) is a derivative of a DNA molecule covered in a), b) or c), per se.

The molecule may also comprise the origin of replication of (I) the 2.5 Md plasmid of L. lactis LL712 deposited under the accession number DSM 5804, or of (II) the 5.2 Md plasmid of L. lactis LL712, or of (III) a plasmid of the same incompatibility group as the 2.5 Md plasmid or the 5.2 Md plasmid of L. lactis LL712.

These DNA molecules are useful for the preparation of the hybrid vectors of the invention or for screening DNA gene libraries or mRNA for further similar DNAs or mRNAs.

Process for the Preparation of the DNA Molecules and of MSP

A further object of the invention is a process for the preparation of a DNA molecule of the invention, i.e. of a hybrid vector or a DNA molecule defined hereinbefore, comprising

A) culturing a host which comprises a DNA molecule of the invention and isolating a DNA molecule of the invention from such a cultured host, or

B) preparing a DNA molecule of the invention by an in vitro synthesis.

The culturing of the hosts is carried out in a conventional nutrient medium which may be supplemented with or deprived of chemical compounds allowing negative or positive selection of the transformants, i.e. such hosts containing the desired DNA molecule together with a selection marker, from the non-transformants, i.e. such hosts lacking the desired DNA molecule.

Any suitable transformable hosts useful in the art may be used, for example suitable bacteria, such as gram-negative bacteria, e.g. E. coli, or gram-positive bacteria, e.g. Bacillus spec. or lactic acid bacteria, such as Lactobacillus spec., or especially Lactococcus spec., e.g. L. lactis, L. lactis diacetylactis or L. cremoris.

Bacteria are transformed by a conventional method and the transformants are identified in conventional manner, e.g. by their resistance, for example against tetracycline.

In particular the described hybrid vectors are propagated in suitable E. coli host strains, such as TG1, HB101 (Maniatis et al (1982)), JM109, MH1 and the like, or in suitable Bacillus strains, or in suitable lactic acid bacteria, e.g. Lactococcus strains such as L. lactis 0230, L. lactis diacetylactis, L. cremoris and the like. The hosts are transformed and selected by methods conventional in the art. The propagated plasmid DNA is isolated from the bacteria by conventional methods, for example as described by Birnboim & Doly (1979).

A DNA molecule of the invention can also be prepared by an in vitro synthesis according to conventional methods. The in vitro synthesis is especially applicable for the preparation of smaller fragments, e.g. of the DNA sequences of the MSP gene or of a related gene coding for the promoter or in particular for the signal peptide.

A DNA molecule of the invention can be obtained from a lactic acid bacterium containing such a DNA molecule, in particular from a genomic library thereof or also via mRNA.

In the following the preparation of a DNA molecule which is the approximately 3.5 kbp EcoRI/Sall L. lactis insert of plasmid pUCRS is described in more detail.

As starting material a genomic library of Lactococcus spec., e.g. L. lactis 0230, can be used which is prepared

according to conventional methods, e.g. by partial digestion of genomic DNA of a *L. lactis* strain, e.g. LM 0230, C2 or LL712, with a restriction enzyme, e.g. Sau3A or MboI, and cloning the high molecular weight DNA fragments in a suitable vector, e.g. the *E. coli* plasmid pUN121 or a lambda phage, e.g. λEMBL3.

Any other strain of lactic acid bacteria producing MSP may also serve as source for the genomic library and likewise other suitable vectors may be used as recipient for the fragments.

In order to successfully screen the genomic library for DNA sequences of the present invention a DNA probe is necessary which hybridizes with such a DNA sequence, e.g. with the MSP structural gene. This can be a synthetic DNA probe if the sequence of e.g. the MSP gene or part thereof is known. As neither the MSP protein nor the MSP gene sequence or part thereof was known prior to the invention, the problems of purification of MSP, of the determination of the N-terminus sequence of MSP and of the preparation of hybridizing DNA probes were solved first.

MSP is the most abundant gene product in the supernatant of cultures of *Lactococcus spec.*, e.g. *L. lactis* as judged after TCA precipitation of the supernatant, SDS-polyacrylamide gel electrophoresis of the precipitated proteins and staining of the gel with Coomassie brilliant blue.

For the purification of MSP any source containing it may be used, for example, the supernatant of a culture of lactic acid bacteria, e.g. *Lactococcus spec* such as *L. lactis*. MSP can be purified from a SDS polyacrylamide gel by cutting out the gel piece comprising the major protein band and by eluting MSP.

Further purification methods such as precipitation with acid, e.g. trichloroacetic acid, salting out, desalting, reprecipitation in form of a different salt, dialysis chromatography, e.g. affinity chromatography, ion exchange chromatography, gel permeation chromatography, electrophoresis, e.g. with SDS-polyacrylamide gel, isoelectric focusing, electroelution, and the like, or any combination thereof can also be applied in order to obtain pure MSP.

In the present invention, a MSP with an apparent molecular weight of about 56 kD was isolated in pure form by precipitating the proteins from the supernatant of *L. lactis* LM0230 with trichloroacetic acid, by a subsequent electrophoresis of the precipitated protein on a SDS polyacrylamide gel, cutting out the region with the major protein band and electroeluting MSP from the gel. This method is also suitable for the isolation of proteins related to MSP having another apparent molecular weight. The amino acid sequence of MSP was partly determined by sequencing and partly deduced from the DNA sequence. It is shown in the sequence listing under SEQ ID No. 1 and extends from position 28 up to 466 of the amino acid sequence depicted.

The pure MSP protein is also a subject of the invention.

The sequencing of the N-terminus of MSP was performed in conventional manner and revealed the following amino acid sequence:

X-X-Asn-Ser-Asp-Ile-Ala-Lys-Gln-Asp-Ala-Thr-Ile-Ser-X-Ala-Gln-Ser-Ala-Lys-Ala-Gln-Ala-Gln-Ala-Gln-Val-Asp. The first two amino acids and the amino acid in position 15 which are indicated with "X" have not been determined.

Based on the sequence of the amino acids in position 5 to 13 the following oligonucleotide mixture was synthesized: GAN₁ ATN₂ GCI AAN₃ CAN₃ GAN₁ GCI AC. In this nucleotide sequence, N₁ is T or C; N₂ is T, C or A; N₃ is A or G. A represents a nucleotide with the base Adenine, T with Thymine, C with Cytidine, G with Guanosine and I with Inosine. The oligonucleotide mixture was radioactively marked in conventional manner and used to screen a genomic library of *Lactococcus spec.*, particularly *L. lactis* LM0230.

DNA probes containing sequences encoding MSP amino acid sequences and having at least about 14 bp can be used for screening for nucleic acids comprising related MSP genes or part thereof, which includes also the screening for pre-MSP.

For screening purposes the DNA probes are radioactively labelled at their 5' end by methods known in the art using ³²P-ATP and T4 kinase. Host microorganisms or phages carrying nucleic acids of the present invention as an insert, are identified by hybridization with the labelled DNA probe on filter replicas of the gene library.

The hybridization conditions used are conventional and may be more or less stringent, e.g. simply by choosing different temperatures.

The hybridizing part of a DNA clone of the library which hybridizes with the oligo-nucleotide mixture was partly sequenced according to conventional methods. The determined sequence comprises nearly the entire functional MSP gene. The sequence is depicted in the sequence listing under SEQ ID No. 1.

The 3.5 kbp EcoRI/Sall *L. lactis* insert can be isolated from positive clones of a genomic library by digestion with EcoRI and Sall and subsequent purification of the insert according to conventional methods, e.g. using agarose gel electrophoresis. Fragments can be prepared according to conventional methods and can be ligated into a suitable vector, e.g. into M13mp18 to generate M13mp18RS or into pUC18 to generate pUCRS.

Likewise, any other *L. lactis* insert comprising MSP sequences, for example larger derivatives, variants or fragments of the 3.5 kbp EcoRI/Sall insert, or DNA molecules which hybridize with the 3.5 kbp EcoRI/Sall insert or comprise a promoter region which is naturally linked to such a hybridizing DNA molecule can be isolated from a genomic library of lactic acid bacteria.

Fragments of the DNA molecule according to a) to c) may be obtained in conventional manner, e.g. by isolation of the fragments after digestion of the insert with suitable exo- or endonucleases, e.g. with exonuclease III, Bal31 or

S1 or restriction endonucleases, such as Sau3A, HindIII, and the like. Fragments may also be obtained by in vitro DNA synthesis according to conventional methods.

Mutants of the DNA molecule covered in either of a) to c) hereinbefore, for example inversion, deletion, insertion or point mutants, can be prepared according to conventional methods, for example in vivo or in vitro by site-directed mutagenesis (see review article of Zoller and Smith 1983, Botstein and Shortle, 1985, or Norris et al., 1983) using mutagenic oligonucleotide primers or by deleting DNA fragments between two restriction sites by cutting with suitable restriction enzymes and religating the DNA, optionally in diluted solution.

In the following the preparation of a recombinant DNA molecule according to the invention comprising a DNA sequence which functions as an origin of replication is described in more detail.

Plasmids are isolated from L. lactis LL712 according to conventional methods, e.g. as described by Birnboim and Doly (1979) with the modifications described in the examples, and are separated in conventional manner, e.g. using chromatography techniques, for example agarose gel chromatography as described in Maniatis et al. (1982). The 2.5 Md plasmid and the 5.2 Md plasmid are isolated and fragmented according to conventional methods.

The fragment mixtures thus obtained are ligated with a suitable vector according to conventional methods; e.g. as described in Maniatis et al. (1982) and these ligation mixtures are used to transform in conventional manner a suitable intermediate host strain.

A suitable vector carries a marker gene for selection in lactic acid bacteria, e.g. a resistance marker gene, for example the erythromycin resistance gene, and an origin of replication which does not function in lactic acid bacteria but in a bacterium suitable as an intermediate host for cloning the fragments of the L. lactis plasmids. A suitable vector comprises also a marker gene for selection in the intermediate host which may be identical with the marker gene which functions in the lactic acid bacterium. A suitable intermediate host is, for example, an E. coli strain, e.g. E. coli TG1, and a suitable cloning vector then is an E. coli vector, for example a pUC18 derivative carrying an erythromycin resistance gene, such as pUC383, the construction of which is described hereinafter in the Examples.

The intermediate host cells which are transformed with a vector comprising fragments of the L. lactis 2.5 Md or 5.2 Md plasmid optionally may be selected in conventional manner which depends on the type of intermediate host cell and the vector used. Selection markers may be, for example, resistance markers or genes encoding a screenable marker enzyme, e.g. the product of the E. coli lacZ gene, β -galactosidase, provided that the host is an E. coli strain defective in the genomic lacZ gene. The selection marker genes may be disturbed by the insertion of a DNA fragment. If the marker is lacZ, the host cells which carry a vector with a fragment inserted into the lacZ gene are no more able to convert X-Gal into a blue dye and, as a consequence, positive clones on X-Gal comprising agar plates remain white after induction of the expression of the lacZ gene with IPTG.

Vectors carrying an insert derived from the L. lactis 2.5 Md or 5.2 Md plasmid are tested for their ability to replicate in a plasmid free L. lactis strain, e.g. in L. lactis 0230. For this purpose the plasmids are isolated from the intermediate host and transformed into L. lactis 0230 cells according to conventional methods, e.g. as described in Powell et al. (1988). Replicating vectors comprise a DNA insert derived from the L. lactis 5.2 Md or 2.5 Md plasmid which functions as an origin of replication. A DNA molecule comprising this function may be isolated from the replicating vectors, fragmented, mutated and the like and can be used to construct recombinant DNA molecules according to the invention, e.g. cloning or expression vectors replicating in lactic acid bacteria. The L. lactis 2.5 Md or 5.2 Md plasmid derived origins of replication are also functional in bacteria other than lactic acid bacteria, e.g. in Bacillus spec. or E. coli. Therefore, vectors comprising such DNA fragments can be used as shuttle vectors.

Fragments of the 2.5 Md or 5.2 Md plasmid carrying an origin of replication can also be identified and isolated according to the method described hereinbefore from a fragment mixture which was obtained by fragmentation of the whole plasmid pool of L. lactis LL712. Which of the cloned fragments carrying an origin of replication is derived from the 2.5 Md or 5.2 Md plasmid may be determined by hybridizing the cloned fragments with the plasmids of the plasmid pool derived from L. lactis LL712 which were separated on an agarose gel or by comparison of the restriction pattern of the cloned fragments and of the plasmids of the pool.

Recombinant DNA molecules according to the invention which comprise the whole DNA sequence of the L. lactis 2.5 Md or 5.2 Md plasmid or a plasmid of the same incompatibility group as the 2.5 Md or 5.2 Md plasmid may be obtained, for example, by cutting the respective plasmid with a suitable restriction endonuclease and ligating it, for example, with a DNA fragment comprising a homologous or heterologous structural gene, a promoter region or vector sequences or with a linker fragment or the like. Suitable restriction endonucleases have only one recognition and cleavage site in the respective plasmid.

A recombinant DNA molecule comprising a sequence with origin of replication function derived from the 2.5 Md or the 5.2 Md plasmid can be obtained, for example, by a method comprising the preparation of the plasmid pool of L. lactis LL712 according to conventional methods, the identification of the 2.5 Md or 5.2 Md plasmid, e.g. by estimating the molecular weight in agarose gel electrophoresis, fragmentation of the respective plasmid, e.g. with suitable restriction enzymes, preparation of a fragment comprising the origin of replication function, e.g. one of the fragments mentioned hereinbefore, and ligating such a fragment or a mixture comprising such fragment with a DNA molecule com-

prising no origin of replication, e.g. a cloning vector depleted of its origin of replication, and selecting for DNA molecules replicating in at least L. lactis and E. coli.

Plasmids comprising an origin of replication of the same incompatibility group as the 2.5 Md or 5.2 Md plasmid can be identified because they cannot be maintained together with the 2.5 Md or 5.2 Md plasmid, respectively, within the same cell.

The invention concerns also the use of a DNA molecule or a recombinant DNA molecule of the invention for the preparation of hybrid vectors for the expression of a structural gene. Examples of such structural genes are given hereinbefore. The hybrid vectors can be prepared according to conventional methods using enzymes such as restriction enzymes, DNA polymerases, DNA ligases and the like.

Example 4 and 5 describe exemplary in more detail the preparation of hybrid vectors for the expression of a structural gene in lactic acid bacteria, particularly in L. lactis, or in Bacillus spec., particularly B. thuringiensis.

Transformed hosts and process for the preparation thereof

The invention concerns further a bacterial host transformed with a hybrid vector of the invention.

Transformed bacterial hosts according to the invention are suitable for the cloning, amplification and/or preparation of a hybrid vector comprising a DNA molecule defined in of a) to e). The hybrid vectors can replicate in such hosts and are not lost under selective pressure in the cell population during proliferation. The host which can be used depends on the origin of replication comprised in the hybrid vector. In the case that the hybrid vector comprises a DNA sequence with origin of replication function according to e), a suitable host is, for example, any strain of E. coli, Bacillus spec. or Lactococcus spec., which does not contain a plasmid with an origin of replication of the same incompatibility group.

In the case that the hybrid vector comprises a homologous or heterologous structural gene fused in a proper reading frame with the DNA sequence coding for the MSP signal peptide, a transformed host of the invention is such which is suitable for the production of a secreted homologous or heterologous protein, e.g. a strain of L. lactis.

An example for a transformed host according to the invention is an E. coli strain, e.g. TG1, C600 (Huynh et al (1985)) or HB101 transformed with pUCRS, pSC12, pSC12ΔP, pSC12ΔN, pSC12ΔNP, pSC18, pSC18ΔN, pSC18ΔP, pVAHIR, pVAHIR-1, pSC12HIR, pSC12HIR-1, or pSC12HIRTerm, or L. cremoris, or a plasmid-free L. lactis, e.g. L. lactis LM0230, or a Bacillus strain, e.g. B. thuringiensis, transformed with either of these plasmids. Preferred are E. coli TG1 transformed with pUCRS, L. lactis transformed with pUCRS, pSC12 HIRTerm, pSC12 or pSC18, and B. thuringiensis transformed with pVAHIR or pVAHIR-1.

The invention concerns also a method for the preparation of such transformants comprising treatment of a host under transforming conditions with a recombinant DNA molecule of the present invention, especially a hybrid vector of the invention, optionally together with a selection marker gene and selecting the transformants.

Process for the preparation of polypeptides

The invention concerns further a method for the preparation of a polypeptide, characterized in that a homologous or heterologous structural gene is fused in proper reading frame with the DNA sequence coding for the MSP signal peptide, that a suitable host such as a gram positive bacterial host, for example a lactococcal host, e.g. L. lactis LM0230, or Bacillus spec., e.g. B. thuringiensis, is transformed with a hybrid vector comprising such fused gene and that the polypeptide encoded by said gene is secreted from the host cells. When required, the polypeptide is isolated from the supernatant according to conventional methods. In a preferred embodiment of the invention a bacterial host is used which does not secrete a protease in its culture medium which degrades the expressed protein.

Such a preferred embodiment of the invention is, for example, the production of proteins, e.g. desulfatohirudin, which are secreted from L. lactis LM0230 cells transformed with a hybrid vector of the invention comprising a gene encoding such protein, e.g. pSC12HIR, pSC12HIR-1 or in particular pSC12HIRTerm. Concentrations of active desulfatohirudin can be obtained from supernatant of stationary phase cultures of L. lactis LM0230 cells transformed with pSC12HIR or pSC12HIR-1. In the supernatant of L. lactis LM0230 cells transformed with pSC12HIRTerm the yield in desulfatohirudin is increased. Surprisingly, the levels of heterologous gene products, e.g. desulfatohirudin, do not decrease after prolonged incubation, e.g. over night, of the stationary phase cultures indicating absence of proteolytic degradation of the heterologous gene products in the supernatant.

In another embodiment of the invention the cells may be collected from the nutrient medium either in the log phase or in the stationary phase of the culture, e.g. by centrifugation or filtration, and resuspended in a smaller volume, e.g. in about 1 % up to 20 % of the original culture volume, of fresh nutrient medium or suitable buffer solution. Incubation of the resuspended cells results in an increased yield of heterologous gene product which is secreted into the supernatant. For example, in a culture of L. lactis LM0230 transformed with pSC12HIR an increased yield in desulfatohirudin can be obtained in the supernatant if the cells of a stationary phase culture are collected by centrifugation, resuspended in about 1/10 vol. of fresh nutrient medium and incubated for about 30 min at about 30°C.

Another embodiment of the invention is the production of MSP which is secreted from a gram positive host cell transformed with a hybrid vector of the invention, e.g. pUCRS, according to the methods described above.

A further embodiment of the invention is the production of desulfatohirudin in *B. thuringiensis*, preferentially in strain HD1 cryB (DSM 4574), transformed with pVAHIR or pVAHIR-1.

Short description of the figures

Figure 1: Physical map of plasmid pUC838. The map positions are approximate and given in kbp. The pUC 18 moiety extends from map position 0 up to 2.7 and the pVA838 moiety carrying the erythromycin resistance gene extends from map position 2.7 up to 4.4. The restriction sites shown in brackets were destroyed during the preparation of the recombinant molecule. The meanings of the abbreviations are: *amp^r*, ampicillin resistance gene; *ery^r*, erythromycin resistance gene; *ori_{puc}*, origin of replication derived from pUC 18.

Figure 2: Physical map of plasmid pSC12. The map positions are approximate and are given in kbp. pUC18 DNA extends from position 0 up to 2.7. The 1.7 kb erythromycin resistance carrying fragment of pVA838 extends from position 2.7 up to 4.4. The *L. lactis* 2.5 Md plasmid derived insert extends from position 4.4 up to 8.0/0. The restriction sites shown in brackets were destroyed during the preparation of the recombinant molecules. The meanings of the abbreviations are: PL1, pUC polylinker region from the EcoRI to the (SmaI) site; PL2, pUC polylinker region from the (SmaI) to the SphI site; *amp^r*, ampicillin resistance gene; *ery^r*, erythromycin resistance gene; *ori_{puc}*, origin of replication derived from pUC; *ori₁₁*, origin of replication derived from lactococcal 2.5 Md plasmid as conducted from deletion analysis.

Figure 3: Physical map of plasmid pSC18. The map positions are approximate and are given in kbp. pUC18 DNA extends from position 0 up to 2.7. The 1.7 kb erythromycin resistance carrying fragment of pVA838 extends from position 2.7 up to 4.4. The *L. lactis* 5.2 Md plasmid derived insert extends from position 4.4 up to 8.5/0. The restriction sites shown in brackets were destroyed during the preparation of the recombinant molecules. The meanings of the abbreviations are: PL1, pUC polylinker region from the EcoRI to the (SmaI) site; PL2, pUC polylinker region from the (BamHI) to the HindIII site; *amp^r*, ampicillin resistance gene; *ery^r*, erythromycin resistance gene; *ori_{puc}*, origin of replication derived from pUC; *ori₁₂*, origin of replication derived from lactococcal 5.2 Md plasmid as conducted from deletion analysis.

Abbreviations

dNTP	deoxy nucleotide triphosphate
HPLC	high performance liquid chromatography
IPTG	isopropyl-β-D-thiogalactopyranoside
kbp	kilo base pairs
kD	kilo Daltons
LB	Luria broth nutrient medium (Gibco/BRL)
Md	Mega Daltons
sCD23	soluble CD23, i.e. 25 k IgE-binding factor
SLPI	Secretory Leucoproteinase Inhibitor

The following examples serve to illustrate present invention, however, they should not be construed as a limitation thereof.

Material and Methods

Strains

L. lactis LM0230 is a plasmidfree derivative of *L. lactis* C2 (Efsthathiou J.D. and L.L McKay, 1977). It is deposited with the Deutsche Sammlung von Mikroorganismen und Zellkulturen (see hereinafter).

L. lactis C2 (NCDO 2031), and *L. lactis* LL712 which latter serve as the source of the 2.5 Md and 5.2 Md plasmids, are closely related or identical (F.L. Davis et al., 1981). The former is obtainable, for example, from the National Collection of Dairy Organisms, UK (NCDO 2031), the latter is deposited with the Deutsche Sammlung von Mikroorganismen und Zellkulturen (see hereinafter).

E. coli TG1 has the genotype K12, Δ(lac-pro), supE, thi, hsdD5/F' traD36, proA⁺B⁺, lacI^q, lacZΔM15 and is obtainable from Amersham; it is described in the manual "Oligonucleotide-directed *in vitro* mutagenesis system" of Amersham.

General Methods:Transformation

E. coli is transformed according to the calcium chloride method as described in Maniatis et al. (1982). Transformation of lactococcal strains is performed by electroporation using a Gene Pulser® apparatus from BioRad and following their protocol for L. lactis LM0230.

Plasmid Preparation

Plasmid DNA from E. coli is prepared according to the method of Birnboim and Doly (1979). The same method is used for the isolation of plasmids from lactococci with the following modifications:

- a) Overnight cultures are grown in M17-G medium (Terzaghi and Sadine, 1975) at 30°C, diluted 1:10 into fresh medium and grown for another two hours.
- b) Incubation with lysozyme is performed also in the presence of mutanolysin (50 mg/l) for 10-20 min at 37°C.

A plasmid pool from L. lactis LL712 is prepared by the same method and the plasmid fraction is further purified by CsCl-ethidium bromide equilibrium centrifugation (Clewell and Helinski, 1969). The three smaller plasmids (1.8 Md, 2.5 Md, 5.2 Md; Gasson, 1983) are enriched in the purified plasmid fraction while the two larger plasmids (9 Md and 33 Md; Gasson, 1983) are present in smaller quantities.

Examples:1. Construction of Shuttle Vectors for Lactic Acid Bacteria and E. coli1.1 Construction of pUC838

pVA838 (ATCC 37160; Macrina et al., 1982) is a lactococcal plasmid which carries the gene for constitutively expressed erythromycin resistance on a 1.7kbp Aval/HindIII restriction fragment. About 10 µg of pVA838 DNA is digested with Aval and HindIII. The ends of the DNA fragments are blunt ended with Kleenow enzyme in the presence of all four dNTPs. The fragments are separated on an agarose gel and the 1.7kbp Aval/HindIII fragment is recovered from the gel by electroelution.

About 200 ng of said fragment and 100 ng of Smal-cut pUC18 (Norranders et al.) are ligated in the presence of T4-ligase as suggested by Rusche et al. and the ligation mixture is used to transform E. coli TG1. White colonies are picked from LB agar plates containing X-Gal, IPTG and 100 mg/l of ampicillin. Correct clones are identified by isolating and analyzing their plasmid DNA, and by their ability to grow on LB agar plates containing erythromycin (100 mg/l). The resulting plasmid is designated pUC838. The restriction sites and characteristics of pUC838 are given in Figure 1.

1.2 Construction of pSC12

The plasmids in the pool isolated from L. lactis LL712 are separated on a preparative agarose gel. The second DNA band from bottom corresponds to the ccc form of the 2.5 Md plasmid. It is cut out and the plasmid is electroeluted. Digestion of this plasmid with SphI and agarose gel electrophoresis reveals a single 3.5 kbp fragment indicating the presence of a single or several closely spaced SphI cleavage sites in the 2.5 Md plasmid.

pUC838 is cut at its unique SphI site and treated with calf intestinal phosphatase. About 50 ng of phosphatase treated vector and 150 ng of the 3.5 kbp SphI fragment of the 2.5 Md plasmid is ligated with T4 ligase and the ligation mixture is used to transform competent E. coli TG1 cells. The transformation mixture is grown at 37°C in 1 ml of LB for 90 min and then transferred to 200 ml of LB containing 100 mg/l of erythromycin and grown at 37°C over night. Plasmid DNA is prepared from the cells of this culture and about 2 µg are used to transform L. lactis LM0230. Transformed L. lactis LM0230 cells are selected at 30°C on M-17G agar plates containing 5 mg/l of erythromycin.

Plasmid DNA is isolated from several clones of the transformed L. lactis LM0230 cells and subjected to restriction enzyme analysis. All transformants investigated contained a 8.0 kbp plasmid, pSC12, from which the 3.5 kbp SphI insert can be recovered. A physical map of pSC12 is given in Figure 2.

pSC12 is shuttled several times between L. lactis and E. coli without any obvious changes in restriction pattern.

1.3 Construction of pSC18

Plasmid pool DNA from *L. lactis* LL712 is partially restricted with Sau3AI by incubating 500 ng aliquots of DNA with different amounts of restriction enzyme. Samples containing partially cut DNA are identified on an agarose gel. This DNA is ligated to pUC838 vector which is previously cut at its unique BamHI site and treated with calf intestinal phosphatase.

Transformation of *E. coli* TG1 plasmid isolation and subsequent transformation of *L. lactis* LM0230 is done as described for pSC12. Again, restriction digests are performed on plasmid DNA isolated from several clones of *L. lactis* LM0230 transformants. They carry a 8.8 kbp plasmid, pSC18. Shuttling the plasmid several times between *L. lactis* and *E. coli* didn't result in any changes in restriction pattern indicating stability of the construct in both cell types. A physical map of pSC18 is given in Fig. 3.

2. Identification of Lactococcal origins of replication

2.1 Deletion Analysis of pSC12 and pSC18

The regions on plasmids pSC12 and pSC18 carrying the lactococcus derived origins of replication are defined by deletion analysis as follows:

DNA between appropriate restriction sites is deleted by cutting the plasmids with the corresponding restriction enzyme and recircularization with T4 ligase. The approximate location of the restriction enzyme cleavage sites in pSC12 and pSC18 are given in Figures 2 and 3, respectively, and in Table 1.

pSC12ΔP and pSC18ΔP are constructed by deleting the DNA between the two PvuII sites in the pUC18 part of the plasmids pSC12 and pSC18, respectively, thus removing the col EI derived origins of replication.

In pSC12ΔN and pSC18ΔN the smaller NdeI-fragment of pSC12 or pSC18, respectively, is removed.

pSC12ΔR lacks an EcoRV fragment in the lactococcal DNA of pSC12.

In pSC12ΔNP DNA between the PvuII site in pUC and the NdeI site in the lactococcal part of pSC12 is removed. In this case the vector DNA is blunted with Klenow enzyme before religation.

A further derivative of pSC18, pSC18ΔH, lacks DNA spanning the HindIII sites in the plasmid.

All the mutated plasmid DNAs are first isolated from *E. coli* TG1 transformants that are resistant to erythromycin. It is then tested whether they can be used to transform *L. lactis* LM0230 to erythromycin resistance. Failure to do so is assumed to indicate that the lactococcal origin of replication was removed or destroyed by the deletion. When erythromycin resistant transformants are obtained, the plasmid is reisolated and its structural integrity reconfirmed by restriction analysis.

For both the 2.5 Md and the 5.2 Md plasmid a region required for replication in streptococcus is identified. From the fact that the ΔP-derivatives could replicate in *E. coli* it is obvious that the cloned lactococcal origins can also function in the gram-negative *E. coli* bacteria.

Results of these experiments are given in Table 1.

Table 1:

Results of the deletion analysis of pSC12 and pSC18				
Plasmid	Deleted region ¹	Deleted fragment ²	Replication of the plasmid in	
			<i>E. coli</i>	<i>L. lactis</i>
pSC12	none		+	+
pSC12ΔP	PvuII (0.1), PvuII (2.5)	2.4	+	+
pSC12ΔN	NdeI (0.2), NdeI (6.2)	2.0	+	+
pSC12ΔR	EcoRV (5.0), EcoRV (7.5)	2.5	+	-
pSC12ΔNP	NdeI (6.2), PvuII (2.5)	4.3	ND ³	+

¹ The deletions extend between two restriction sites. The approximate positions of the restriction sites in pSC12 or pSC18 are given in the brackets as the distances from the HindIII site in the pUC838-derived moiety of the plasmids which is defined as position (0). They relate to the base positions given in Figures 2 and 3, respectively.

² Approximate length of the deleted fragment in kbp.

³ Not determined.

Table 1: (continued)

Results of the deletion analysis of pSC12 and pSC18				
Plasmid	Deleted region ¹	Deleted fragment ²	Replication of the plasmid in	
			<u>E. coli</u>	<u>L. lactis</u>
pSC18	none		+	+
pSC18ΔP	PvuII (0.1), PvuII (2.5)	2.4	+	+
pSC18ΔN	NdeI (0.2), NdeI (7.5)	1.2	+	+
pSC18ΔH	HindIII (0), HindIII (6.5)	2.0	+	-

¹The deletions extend between two restriction sites. The approximate positions of the restriction sites in pSC12 or pSC18 are given in the brackets as the distances from the HindIII site in the pUC838-derived moiety of the plasmids which is defined as position (0). They relate to the base positions given in Figures 2 and 3, respectively.

²Approximate length of the deleted fragment in kbp.

2.2 Source of the lactococcal origins of replication

The origin of replication in pSC12 is derived from the 2.5 Md plasmid of *L. lactis* LL712 as is clear from the isolation protocol for the DNA fragment carrying the origin.

The restriction enzyme cleavage pattern indicates that the origin of replication in pSC18 stems from the 5.2 Md plasmid.

2.3 Replication of pSC12 and pSC18 in other Lactic Lactococci

Strains of *L. lactis*, *L. lactis diacetylactis* and *L. cremoris* are successfully transformed with both plasmids. The plasmids remain stably associated with the cell population under selective pressure.

3. Cloning of the Gene encoding the Major Secretion Product (MSP) of *L. Lactis* LM 0230

3.1. Isolation of the Major Secreted Protein (MSP) of *L. Lactis* LM 0230

1.5 l of an overnight culture of *L. lactis* LM 0230, grown at 30 °C in M-17G, is centrifuged in a Sorvall GS-3 rotor at 7000 rpm for 20 min at 4 °C. The supernatant is collected and proteins precipitated by adding an equal volume of ice cold 10 % trichloroacetic acid and incubating for 30 min at room temperature. The precipitate is collected by centrifugation in a Sorvall GS-3 rotor at 8 000 rpm and 4 °C for 30 min. The protein pellets are drained and redissolved in 3 ml of SDS-sample buffer (Laemmli, 1970). The sample is neutralized by adding a small amount of 4N NaOH and then dialyzed for 4 h against 4 l of 25 mM Tris-HCl pH 6.8, 0.02 % SDS. The sample recovered has a volume of about 4 ml. 3x conc. SDS-sample buffer is added to the sample in order to enlarge the volume to 6 ml.

Proteins are separated by running 2 ml aliquots on preparative 8 % SDS polyacrylamide gels (Laemmli, 1970) using a Protean cell from BioRad. Small strips are cut alongside from the gel, stained with Coomassie Brilliant Blue and destained in an aqueous solution containing 10 % of Methanol and 10 % of Acetic acid. They serve as markers to identify and cut out the major protein band with an apparent molecular weight of about 56 kD which is cut out of the gel.

The MSP protein is recovered from the gel by electroelution at 150 V for 2 h using a Biotrap apparatus and a buffer containing 20 mM ammonium acetate and 0.01 % SDS. The eluate is dialyzed for 48 h against two changes of 20 mM ammonium acetate, 0.005 % SDS. A sample of the dialyzed protein is run on a 8 % SDS-PAGE and a single sharp protein band is observed with an apparent molecular weight of about 56 kD.

3.2 Analysis of the amino-terminal Sequence of MSP

Amino-terminal sequencing of the isolated MSP protein is performed according to conventional methods using a gas phase sequencer (Applied Biosystems Inc., Model 470A) with HPLC quantitation of phenylthiohydantoin-derivatives of cleaved amino acid residues.

The 28 amino acid long sequence obtained is X-X-Asn-Ser-Asp-Ile-Ala-Lys-Gln-Asp-Ala-Thr-Ile-Ser-X-Ala-Gln-Ser-Ala-Lys-Ala-Gln-Ala-Gln-Ala-Gln-Val-Asp. The first two amino acids and the amino acid in position 15 which are indicated with "X" are not determined.

3.3 Synthesis of a mixed Oligonucleotide Probe

Based on the amino acid sequence Asp-Ile-Ala-Lys-Gln-Asp-Ala-Thr-Ile, which corresponds to the amino acids in position 5 to 13 in the N-terminus sequence of MSP, a mixed oligonucleotide is designed and constructed for the screening of a DNA library. Inosine is inserted at the two positions in the 23-mer where the degeneracy of the genetic code would require all 4 dNTPs in order to decrease the complexity of the probe. The nucleotide sequence of the oligonucleotide mixture is GAN₁ ATN₂ GCI AAN₃ CAN₃ GAN₁ GCI AC. N₁ is T or C; N₂ is T, C or A; N₃ is A or G. A represents a nucleotide with the base Adenine, T with Thymine, C with Cytidine, G with Guanosine and I with Inosine.

3.4. Construction of a Genomic Library of *L. lactis* LM 0230

Chromosomal DNA is isolated from *L. lactis* LM 0230 by a modification of the protocol used for plasmid isolation: after treatment with lysozyme and mutanolysin the cells are incubated for two hours at 56°C with proteinase K (100 mg/l) in 10 mM Tris-HCl pH 8, 20 mM EDTA, 0.5 % SDS. The DNA is extracted once with phenol/chloroform (1 vol:1 vol) and then purified by CsCl density gradient centrifugation.

Chromosomal DNA is partially digested with Sau3AI and size fractionated on sucrose gradients as described in Maniatis. Fragments between about 10-20 kbp are collected.

λ EMBL 3 DNA is cleaved with BamHI and EcoRI and purified by phenol extraction and ethanol precipitation. It is then ligated with the 10-20 kbp fragments in the presence of hexamminecobalt(III)chloride to favour the formation of concatamers. (Rusche J.R. et al., 1985). The ligation mixture is packaged *in vitro* using the Gigapack Plus® system from Stratagene.

Recombinants are selected by plating the library on *E. coli* Q 359 (Kahn J. et al., 1980). A total of about 600 000 recombinant phages is obtained.

3.5 Screening the Library for DNA sequences encoding MSP

About 15 000 recombinant λ EMBL 3 plaques are plated per petri dish (15 cm diameter) and transferred to Plaque-Screen® membranes (NEN). The filters are screened by hybridization with the mixed oligonucleotide probe described in Example 3.3 which was labeled with [³²P]ATP using T4 Kinase according to standard procedures described in Maniatis et al. (1982).

Positive plaques are identified and subjected to a second round of screening at low plaque density.

DNA is prepared from positive phages, digested with restriction enzymes and subjected to Southern analysis as described in Maniatis. Southern blots are probed with the mixed oligonucleotide. It hybridizes to a 3.5 kbp EcoRI/Sall fragment in the *L. lactis* insert of a positive phage. Further probing showed a 2.1 kbp HindIII fragment within the 3.5 kbp EcoRI/Sall fragment to hybridize with the oligonucleotide mixture.

The approximate distances of different restriction sites from the EcoRI cut end of the 3.5 kbp EcoRI/Sall fragment are determined by agarose gel electrophoresis after digestion of the fragment with the respective enzymes and suitable enzyme mixtures. The EcoRI cut end is located about 0.5 kbp away from a HindIII site, about 2.6 kbp from a second HindIII site, and about 3.15 kbp from a third HindIII site.

3.6 Construction of pUCRS

DNA from a positive phage is digested with EcoRI and Sall. The DNA fragments are separated on an 0.6 % agarose gel and the 3.5 kbp EcoRI/Sall fragment is cut out and isolated by electroelution. A three fold molar excess of this fragment is ligated to pUC19 which is cut with EcoRI and Sall and treated with alkaline phosphatase.

The ligation mixture is used to transform *E. coli* TG1 and transformants are selected on plates containing 100 mg/l of ampicillin. Plasmid DNA is prepared from transformants and restriction analyses is performed to identify the correct construct which is designated pUCRS.

E. coli TG1 transformed with pUCRS is deposited with the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ).

3.7. Sequencing the gene for the Major Secretion Product (MSP)

Sequencing is done by the chain termination method (Sanger F. et al., 1977) using Sequenase® from USBC.

The 3.5 kbp EcoRI/Sall fragment prepared as described in Example 3.6 is subcloned into M13 mp18. The resulting plasmid is named M13 mp18RS. The 2.1 kbp HindIII fragment of the 3.5 kbp EcoRI/Sall fragment is cloned into M13 mp19, the resulting plasmid is M13 mp19H. A set of unidirectional deletions is generated from the Sall end of the insert in M13 mp18RS by digesting the DNA with exonucleaseIII and SI (Henikoff, 1984; Yanisch-Perron et al., 1985). The

deletions extend about 2 kbp into the insert. Sequencing these deletions from the universal mp18 primer yielded the sequence information for one strand. Based on this sequence, a set of synthetic oligonucleotides is synthesized and serves as primers to sequence the second strand. Additional sequence information is gained by sequencing exoIII-S1 deletions generated from M13 mp19H with the same method. The DNA sequence of a 1920 bp fragment of the 3.5 kbp EcoRI/Sall fragment is depicted in the sequence listing under SEQ ID No. 1. It comprises part of the MSP promoter region, the DNA sequence encoding the MSP signal peptide and the structural gene for mature MSP.

Analysis of the about 2 kbp long DNA sequence revealed the presence of a single open reading frame (ORF) coding for a protein having 461 amino acid residues. Comparison of the deduced amino acid sequence with the data obtained from sequencing the N-terminus of the isolated MSP identified a mature form of the protein with 434 amino acids which is preceded in frame by 27 amino acids. The peptide formed by the latter amino acids has the structure of a typical signal peptide. Charged residues are found at the N-terminus while the rest of the sequence consists mainly of hydrophobic amino acids. The amino acid composition around the cleavage site follows the rules deduced from other signal peptides (Von Heijne, 1983).

4. Production of recombinant Desulfatohirudin in *Lactococcus lactis*

4.1 Construction of Desulfatohirudin Secretion Plasmids

Plasmid M13 mp19H contains a 2.1 kbp HindIII fragment encoding the N-terminus half of MSP including the signal peptide and about 1.5 kbp of upstream DNA. This plasmid is cleaved at its unique Scal site 129 bp downstream of the COOH-terminus of the signal peptide. A blunt ended 211 bp DNA fragment encoding the genetic information for desulfatohirudin is isolated from plasmid pML310 (published in the European Patent Application EP-A-168 342) and inserted at the cleaved Scal site of M13 mp19H. An in frame fusion between the DNA sequence encoding the MSP signal peptide and the desulfatohirudin structural gene is created by removing the excess DNA separating the codon for the last amino acid of the signal peptide (Ala) and the codon for the first amino acid of desulfatohirudin (Val). This is achieved by oligonucleotide-directed *in vitro* mutagenesis using the system from Amersham. The 29 bp oligonucleotide comprises the last 14 bp of the signal peptide and the first 15 bp of the hirudin gene. The sequence of the resulting fusion is depicted in the sequence listing under SEQ ID No. 2.

The HindIII fragment carrying the fusion is cut out and inserted into the unique HindIII of pSC12. Both orientations of the insert are recovered. The resulting plasmids are named pSC12HIR and pSC12HIR-1, respectively.

Another plasmid (pSC12HIRTerm) is constructed by inserting the *trpA* transcription terminator of *E. coli* (Pharmacia) at the unique HpaI site of pSC12HIR about 125 bp downstream of the desulfatohirudin gene.

4.2 Secretion of Desulfatohirudin by *L. lactis*

The desulfatohirudin fusion plasmids are transformed into *L. lactis* LM 0230. Transformants are grown in M-17G medium supplemented with 2 % of glucose in the presence of erythromycin (5 mg/l) at 30°C. When the cultures reach stationary phase, 1.5 ml samples are centrifuged in an Eppendorf® tube. The culture supernatants are removed and frozen at -70°C. The cultures are then further grown overnight before a second sample is taken, centrifuged, the supernatant removed and also frozen. The production of secreted desulfatohirudin is determined with a bioassay.

The bioassay measures the thrombin inhibition activity of desulfatohirudin in the supernatants collected. A detailed protocol of the bioassay, which is a thrombin inhibition assay is presented hereinafter:

The buffer used for making all dilutions of samples and reagents is 0.2M Tris-HCl, pH 7.5 containing 1.0M NaCl and 0.01 % bovine serum albumin. Thrombin is from human plasma (Protogen AG, Läufelfingen, product No. 80-13-1102), the chromogenic thrombin substrate is Chromozym TH (Boehringer, Mannheim, product No 206849). The p-nitroaniline released from Chromozym TH is measured with a Dynatech MR 600 microplate reader. All assays are performed in microtiter plates (Nunc, MicroWell plates). Each well receives: 50 µl buffer, 50 µl supernatant with unknown concentration of thrombin inhibitory activity, i.e. desulfatohirudin, and 25 µl thrombin solution. The reaction is started by adding 150 µl substrate solution (330 µg/ml) and the plates are incubated for two hours at 37°C. The concentration of thrombin is adjusted to give an $A_{405nm} = 0.8 \pm 0.2$ in an uninhibited control well. Both substrate and thrombin solutions are kept frozen at -20°C and are thawed immediately before being used. A standard curve with known concentrations of a [Tyr⁶³]desulfatohirudin (40, 20, 10, 5, 2.5 and 1.25 ng/ml) is used to convert the OD-measurements into the concentrations of active desulfatohirudin.

In the supernatant of stationary phase cultures of *L. lactis* LM0230 cells transformed with either pSC12HIR or pSC12HIR-1 the same level of desulfatohirudin activity is determined. In the supernatant of a stationary phase culture of *L. lactis* LM0230 cells transformed with pSC12HIRTerm, the level of desulfatohirudin activity is increased by about 50 %. Supernatants of cells transformed as a control with pSC12 contain no desulfatohirudin activity.

The levels of desulfatohirudin in the supernatant did not decrease after prolonged incubation for about 16 h of the

stationary phase cultures. These results indicate the absence of proteolytic degradation in the supernatant of *L. lactis* LM0230.

In one experiment, cells are first collected by centrifugation and then resuspended in 1/10 volume of fresh medium and incubating at 30°C for 30 min expression period. Bioassay measures an about six fold higher level of desulfatohirudin activity than in the experiments described hereinabove. In a more sophisticated form, concentration of the cells for an expression period is a way to increase the concentration of secreted product in the supernatant.

5. Production of recombinant Desulfatohirudin in *Bacillus thuringiensis*

5.1 Construction of plasmid pVAHIR

Plasmid pVA838 (Macrina et al, 1982) is digested with HindIII and the 5.0 kbp fragment carrying the gene for erythromycin resistance and the gram positive origin of replication is isolated from a 0.5 % agarose gel by electroelution.

The HindIII fragment containing the desulfatohirudin gene is isolated in a similar manner from pSC12HIR.

The two fragments are ligated in the presence of T4-ligase and the ligation mixture is directly used to transform *L. lactis* LM 0230 by electroporation. Erythromycin resistant transformants are selected and plasmid DNA is prepared from them.

Restriction analyses of these plasmids is performed to identify correct constructions. Both orientations of the HindIII-desulfatohirudin expression cassette are obtained and can be used for the production of Desulfatohirudin in *Bacillus thuringiensis*. The plasmids are named pVAHIR and pVAHIR-1.

5.2 Secretion of Desulfatohirudin by *B. thuringiensis*

B. thuringiensis strain HID1 cryB (DSM 4574) is transformed with pVAHIR using electroporation (W. Schurter et al, 1989).

Transformants are selected on LB plates containing 20 µg/ml of erythromycin at 27°C or 30°C. Plasmid DNA is prepared from individual transformants by the same method as used for *L. lactis*. Restriction analyses reconfirms that the structure of the isolated plasmids is identical to pVAHIR.

Transformants harbouring pVAHIR or pVA838 - as a control - are grown in LB containing 20 mg/l of erythromycin at 27°C over night. The cultures are diluted 1:200 into fresh medium and grown further at 30°C. The cultures are centrifuged and the supernatants are assayed for desulfatohirudin activity 7 h after dilution.

Desulfatohirudin is detected in the supernatant of cells transformed with pVAHIR, while none is measured in the control with pVA838.

Deposited Microorganisms

The following microorganisms are deposited according to the Budapest Treaty with the Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSM), Mascheroder Weg 1b, D-3300 Braunschweig:

Microorganism	Deposition No.	Date of deposition
<i>E. coli</i> K12 TG1/pUCRS	DSM 5803	February 16, 1990
<i>Lactococcus lactis</i> LL712	DSM 5804	February 16, 1990
<i>Lactococcus lactis</i> LM0230	DSM 5805	February 16, 1990

References:

- Bates, E.E.M. et al. (1989) *Applied and Environmental Microbiology* **55**, 2095.
 Bernton, W.D., and Davis, R.W. (1977), *Science* **196** 180.
 Birnboim, H.C. and Doly, J. (1979) *Nucleic Acids Res.*, **7**, 1513.
 Botstein, D. and Shortle, D. (1985) *Science* **229**, 1193.
 Chassy, B.M. (1987) *FEMS Microbiol. Rev.* **46**, 297.
 Clewell, D. and Helinski, D.R. (1969) *Proc. Natl. Acad. Sci. USA*, **62**, 1159.
 Davies, F.L., Underwood, H.M. and Gasson, J.M. (1981) *J. of Applied Bacteriology*, **51**, 325.
 De Vos, W.M. (1987) *FEMS Microbiol. Rev.* **46**, 281.
 Efsthathiou, J.D. and McKay, L.L. (1977) *J. Bacteriol.*, **130**, 257.
 Gasson, J.M. (1983) *J. Bacteriol.*, **154**, 1.
 Gasson, J.M. and Anderson P.H. (1985) *FEMS Microbiol. Lett.*, **30**, 193.

- Henikoff, S. (1984) *Gene*, 28, 351.
 Huynh et al, in: *DNA Cloning* 1:56-110, D.M. Glover ed., IRL Press Oxford (1985)
 Jos, M. (1985) *Applied and Environmental Microbiology* 50, 540.
 Karn, J., Brenner, S., Barnett, L., and Cesareni, G. (1980) *Proc. Natl. Acad. Sci. USA*, 77, 5172.
 5 Laemmli, U.K. (1970) *Nature*, 227, 680.
 Macrina, F.L. et al., in *Genetic Engineering of Microorganisms for Chemicals*, Plenum Press, New York, (1982).
 Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular cloning, a Laboratory Manual*. Cold Spring Harbor Laboratory, New York.
 Norrander, J. et al. (1983) *Gene*, 26, 101.
 10 Norris, K. et al. (1983) *Nucl. Acids Res.* 11, 5103.
 Powell et al. (1988) *Applied and Environmental Microbiology* 54, 6.
 Oto, R. et al. (1982) *Applied and Environmental Microbiology* 43, 1272.
 Rusche, J.R. and Howards-Flanders, P. (1985) *Nucleic Acids Res.*, 13, 1997.
 Sanger, F., Milken, S., Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA*, 74, 5463.
 15 Schurter, W., Geiser, M. and Mathe, D. (1989) *Mol. Gen. Genet.* 218, 177.
 Terzaghi, B.K. and Sadine, N.E. (1975) *Applied Microbiology*, 29, 807.
 Von Heijne, G. (1983) *Eur. J. Biochem.*, 133, 17.
 Yanisch-Perron, C., Vieira, J. and Messing, J. (1985) *Gene*, 33, 103.
 20 Zoller, M.J. and Smith, M. (1983) *Methods Enzymol.* 100, 468.

Sequence Listing

SEQ ID No. 1

- 25 SEQUENCE TYPE: Nucleotide with corresponding protein
 SEQUENCE LENGTH: 1920 base pairs
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: genomic DNA
 30 ORIGINAL ORGANISM SOURCE: *Lactococcus lactis* LM0230 (DSM 5805)
 IMMEDIATE EXPERIMENTAL SOURCE: Plasmid pUCRS (DSM 5803)
 POSITION IN THE GENOME: chromosomal
 FEATURES: from

- 35 1 to 410 part of the promoter region
 411 to 491 MSP signal peptide
 492 to 1793 mature MSP protein

- 40 PROPERTIES: Major secretion product (MSP) gene of *L. lactis*.

TTTAGGTATT TACGGAATTG CGACCTTATT GTTCCCACTT	40
ATTGCTCTTT TTGTATATAA TATACAAATA ACTATATTTA	80
CTAATCGCTG GACAAGGCTT TTTACAACAA TTATTATTGT	120
GACCGCTTTT GAAGTTTTTA GTGCAATCAT TATGACAGCT	160
50 TTTGGATTTG CCCAACTTCA GTTTATCAAA TTTGTTGTTT	200
ACCAGTTAGC GCCTACACTT TTGCTCAATA TTATCTTAGC	240
55 TGTAGCCTTA CAATTCCCTT TAGAAATCTT TTACAGATTA	280
AAGAAAAGTC ATGTAAGATA CAATTAGAAA GTGTTTTGTA	320

EP 0 449 770 B1

	ATCATAAAGA AATATTAAGG TGGGGTAGGA ATAGTATAAT	360
5	ATGTTTATTC AACCGAACTT AATGGGAGGA AAAATTAAAA	400
	AAGAACAGTT ATG AAA AAA AAG ATT ATC TCA GCT Met Lys Lys Lys Ile Ile Ser Ala	434
10	5	
	ATT TTA ATG TCT ACA GTG ATA CTT TCT GCT GCA Ile Leu Met Ser Thr Val Ile Leu Ser Ala Ala	467
15	10 15	
	GCC CCG TTG TCA GGT GTT TAC GCT GAC ACA AAC Ala Pro Leu Ser Gly Val Tyr Ala Asp Thr Asn	500
20	20 25 30	
	TCA GAT ATT GCT AAA CAA GAT GCG ACA ATT TCA Ser Asp Ile Ala Lys Gln Asp Ala Thr Ile Ser	533
25	35 40	
	AGC GCG CAA TCT GCT AAA GCA CAA GCA CAA GCA Ser Ala Gln Ser Ala Lys Ala Gln Ala Gln Ala	566
30	45 50	
	CAA GTT GAT AGC TTG CAA TCA AAA GTT GAC AGC Gln Val Asp Ser Leu Gln Ser Lys Val Asp Ser	599
35	55 60	
	TTA CAA CAA AAG CAA ACA AGT ACT AAA GCA CAA Leu Gln Gln Lys Gln Thr Ser Thr Lys Ala Gln	632
40	65 70	
	ATC GCT AAA ATC GAA AGC GAA CGT AAA GCA CTT Ile Ala Lys Ile Glu Ser Glu Arg Lys Ala Leu	665
45	75 80 85	
	AAT GCT CAA ATT GCT ACT TTG AAC GAA AGT ATC Asn Ala Gln Ile Ala Thr Leu Asn Glu Ser Ile	698
50	90 100	
	AAA GAA CGT ACA AAG ACA TTG GAA GCT CAA GCA Lys Glu Arg Thr Lys Thr Leu Glu Ala Gln Ala	731
55	105 110	

EP 0 449 770 B1

	CGT AGT GCT CAA GTT AAC AGC TCA GCA ACA AAT	764
	Arg Ser Ala Gln Val Asn Ser Ser Ala Thr Asn	
5	115 120	
	TAT ATG GAT GCT GTT GTT AAT TCA AAA TCT TTG	797
	Tyr Met Asp Ala Val Val Asn Ser Lys Ser Leu	
10	125 130	
	ACA GAT GTT ATT CAA AAA GTA ACA GCT ATT GCT	830
	Thr Asp Val Ile Gln Lys Val Thr Ala Ile Ala	
15	135 140 145	
	ACT GTT TCT AGT GCC AAC AAA CAA ATG TTG GAA	863
	Thr Val Ser Ser Ala Asn Lys Gln Met Leu Glu	
20	150 155	
	CAA CAA GAA AAA GAG CAA AAA GAG CTT AGC CAA	896
	Gln Gln Glu Lys Glu Gln Lys Glu Leu Ser Gln	
25	160 165	
	AAG TCA GAA ACT GTT AAA AAG AAC TAC AAC CAG	929
	Lys Ser Glu Thr Val Lys Lys Asn Tyr Asn Gln	
30	170 175	
	TTC GTT TCT CTT TCA CAA AGT TTG GAT TCT CAA	962
	Phe Val Ser Leu Ser Gln Ser Leu Asp Ser Gln	
35	180 185	
	GCT CAA GAA TTG ACT TCA CAA CAA GCT GAA CTC	995
	Ala Gln Glu Leu Thr Ser Gln Gln Ala Glu Leu	
40	190 195 200	
	AAA GTT GCG ACT TTG AAC TAT CAA GCA ACA ATT	1028
	Lys Val Ala Thr Leu Asn Tyr Gln Ala Thr Ile	
45	205 210	
	GCA ACT GCG CAA GAT AAA AAA CAA GCT TTA TTA	1061
	Ala Thr Ala Gln Asp Lys Lys Gln Ala Leu Leu	
50	215 220	
55		

EP 0 449 770 B1

	GAT GAA AAA GCA GCT GCA GAA AAA GCA GCT CAA	1094
	Asp Glu Lys Ala Ala Ala Glu Lys Ala Ala Gln	
5	225 230	
	GAA GCA GCT AAA AAA CAA GCG GCT TAT GAA GCT	1127
	Glu Ala Ala Lys Lys Gln Ala Ala Tyr Glu Ala	
10	235 240	
	CAA CAA AAA GAA GCA GCA CAA GCA CAA GCA GCT	1160
	Gln Gln Lys Glu Ala Ala Gln Ala Gln Ala Ala	
15	245 250 255	
	TCA ACA GCA GCA ACT GCT AAA GCT GTA GAA GCA	1193
	Ser Thr Ala Ala Thr Ala Lys Ala Val Glu Ala	
20	260 265	
	GCA ACT TCA TCA GCT TCT GCT TCA TCT AGT CAA	1226
	Ala Thr Ser Ser Ala Ser Ala Ser Ser Ser Gln	
25	270 275	
	GCT CCA CAA GTA AGT ACA AGC ACT GAT AAT ACA	1259
	Ala Pro Gln Val Ser Thr Ser Thr Asp Asn Thr	
30	280 285	
	ACA TCA AAT GCT AGT GCC TCA AAC AGT TCT AAT	1292
	Thr Ser Asn Ala Ser Ala Ser Asn Ser Ser Asn	
35	290 295	
	AGT TCA TCA AAC TCA AGT TCA AGT TCT AGC AGT	1325
	Ser Ser Ser Asn Ser Ser Ser Ser Ser Ser Ser	
40	300 305 310	
	TCA TCA AGC TCA AGC TCA AGC TCA AGT AAT TCT	1358
	Ser Ser Ser Ser Ser Ser Ser Ser Ser Asn Ser	
45	315 320	
	AAT GCT GGT GGG AAT ACA AAT TCA GGC ACT AGT	1391
	Asn Ala Gly Gly Asn Thr Asn Ser Gly Thr Ser	
50	325 330	
	ACT GGA AAT ACT GGA GGA ACA ACT ACT GGT GGT	1424
	Thr Gly Asn Thr Gly Gly Thr Thr Thr Gly Gly	
55	335 340	

EP 0 449 770 B1

	AGC GGT ATA AAT AGT TCA CCA ATT GGA AAT CCT	1457
	Ser Gly Ile Asn Ser Ser Pro Ile Gly Asn Pro	
5	345 350	
	TAT GCT GTT GGT GGA TGT ACT GAC TAT GTA TGG	1490
	Tyr Ala Val Gly Gly Cys Thr Asp Tyr Val Trp	
10	355 360 365	
	CAA TAC TTT GCT GCA CAA GGA ATT TAT ATC AGA	1523
	Gln Tyr Phe Ala Ala Gln Gly Ile Tyr Ile Arg	
15	370 375	
	AAT ATC ATG CCT GGT AAT GGT GGA CAA TGG GCT	1556
	Asn Ile Met Pro Gly Asn Gly Gly Gln Trp Ala	
20	380 385	
	TCT AAT GGA CCT GCC CAA GGC GTG CTC CAT GTT	1589
	Ser Asn Gly Pro Ala Gln Gly Val Leu His Val	
25	390 395	
	GTA GGA GCT GCT CCT GGT GTT ATC GCA TCA AGC	1622
	Val Gly Ala Ala Pro Gly Val Ile Ala Ser Ser	
30	400 405	
	TTC TCA GCT GAT TTT GTT GGA TAT GCA AAC TCA	1655
	Phe Ser Ala Asp Phe Val Gly Tyr Ala Asn Ser	
35	410 415 420	
	CCT TAC GGT CAC GTA GCT ATT GTA AAA TCA GTT	1688
	Pro Tyr Gly His Val Ala Ile Val Lys Ser Val	
40	425 430	
	AAT TCA GAT GGT ACA ATT ACT ATC AAA GAA GGC	1721
	Asn Ser Asp Gly Thr Ile Thr Ile Lys Glu Gly	
45	435 440	
	GGA TAT GGT ACA ACT TGG TGG GGA CAT GAA CGT	1754
	Gly Tyr Gly Thr Thr Trp Trp Gly His Glu Arg	
50	445 450	
	ACT GTA AGT GCG TCT GGT GTT ACT TTC TTG ATG	1787
	Thr Val Ser Ala Ser Gly Val Thr Phe Leu Met	
55	455 460	

EP 0 449 770 B1

CCA AAC TAG AAAAAAGTCT TAATAAATAA AAAATAGTGG 1826
Pro Asn

5 465

TTTGATAGTG GGGAATAATT TTCCTTCTGT CAAATCATT 1866

10 TTTATTATTG TGGTATAATA ATAAGGAAAA ATGATAAGGG 1906

GATAGATACA AATG 1920

SEQ ID No.: 2

15

SEQUENCE TYPE: Nucleotide with corresponding polypeptide

SEQUENCE LENGTH: 279 bp

STRANDEDNESS: double

TOPOLOGY: linear

20

MOLECULE TYPE: recombinant

ORIGINAL SOURCE

ORGANISMS: *L. Lactis* LM0230 (DSM 5805), *Hirudo medicinalis*

IMMEDIATE EXPERIMENTAL SOURCE: Plasmid pUCRS (DSM 5803), plasmid pML310 (see European patent application EP-A-168 342)

25

FEATURES: from

1 to 81 bp MSP signal peptide

82 to 279 bp coding region for desulfatohirudin

30

PROPERTIES: Fusion of the *L. Lactis* LM0230 DNA encoding the MSP signal peptide and of the hirudin structural gene for the production of secreted hirudin in bacteria.

35

ATG AAA AAA AAG ATT ATC TCA GCT ATT TTA ATG TCT 36

Met Lys Lys Lys Ile Ile Ser Ala Ile Leu Met Ser

-25

-20

40

ACA GTG ATA CTT TCT GCT GCA GCC CCG TTG TCA GGT 72

Thr Val Ile Leu Ser Ala Ala Ala Pro Leu Ser Gly

-15

-10

-5

45

GTT TAC GCT GTT GTT TAC ACC GAC TGC ACC GAA TCT 108

Val Tyr Ala Val Val Tyr Thr Asp Cys Thr Glu Ser

50

1

5

55

	GGT CAG AAC CTG TGC CTG TGC GAA GGT TCT AAC GTT	144
5	Gly Gln Asn Leu Cys Leu Cys Glu Gly Ser Asn Val	
	10 15 20	
	TGC GGT CAG GGT AAC AAA TGC ATC CTG GGT TCT GAC	180
10	Cys Gly Gln Gly Asn Lys Cys Ile Leu Gly Ser Asp	
	25 30	
	GGT GAA AAA AAC CAG TGC GTT ACC GGC GAA GGT ACC	216
15	Gly Glu Lys Asn Gln Cys Val Thr Gly Glu Gly Thr	
	35 40 45	
	CCG AAA CCG CAG TCT CAC AAC GAC GGT GAC TTC GAA	252
20	Pro Lys Pro Gln Ser His Asn Asp Gly Asp Phe Glu	
	50 55	
25	GAA ATC CCG GAA GAA TAC CTG CAG TAG	279
	Glu Ile Pro Glu Gln Tyr Leu Gln	
30	60 65	

Claims

- 35
1. hybrid vector comprising
 - a) the approximately 3.5 kbp EcoRI/Sall *L. lactis* insert of the plasmid pUCRS deposited under the accession number DSM 5803, or a functional fragment thereof retaining promoter, signal or structural functions, or
 - 40 b) a DNA sequence which hybridizes with said insert or with a functional fragment thereof, or such a hybridizing DNA sequence operatively linked to its natural promoter region or
 - 45 c) a degenerate sequence of the DNA sequence covered in a) which encodes a signal peptide,
 2. A hybrid vector according to claim 1 comprising the origin of replication of a plasmid belonging to the same incompatibility group as the 2.5 Md or the 5.2 Md plasmid of *L. lactis* strain LL712 deposited under the accession number DSM 5804, which origin can be obtained from the approximately 1.8 kbp long NdeI/SphI fragment of the 2.5 Md plasmid of *L. lactis* LL712 or the 1.0 kbp NdeI/HindIII fragment of the 5.2 Md plasmid of *L. lactis* LL712.
 - 50 3. A hybrid vector according to claim 1 or 2 comprising the approximately 3.5 kbp EcoRI/Sall *L. lactis* insert of the plasmid pUCRS, deposited under the accession number DSM 5803, or a functional fragment thereof.
 4. A hybrid vector according to claim 1 or 2 in which said functional fragment retains the promoter functions of the major secretory protein gene.
 - 55 5. A hybrid vector according to claim 4 in which the functional fragment retaining promoter function spans the region of 1500 bp upstream of the DNA sequence encoding the signal peptide.

6. A hybrid vector according to claim 5 in which the functional fragment retaining promoter function spans a region between 100 bp up to 1000 bp upstream of the DNA sequence encoding the signal peptide.
- 5 7. A hybrid vector according to claim 1 or 2 in which said functional fragment retains the function of the major secretory protein signal sequence.
8. A hybrid vector according to claim 7 in which the functional fragment retaining the signal function encodes the 27 amino acid long signal sequence disclosed in SEQ ID No. 2.
- 10 9. A hybrid vector according to claim 7 in which the functional fragment retaining the signal function extends from a base in about position 411 of the DNA sequence with SEQ ID No. 1 up to about position 491.
10. A hybrid vector according to claim 3 in which said functional fragment retains the promoter and the signal functions of the major secretory protein gene.
- 15 11. A hybrid vector according to claim 10 in which the functional fragment is selected from the group of fragments extending from any base between the EcoRI cut end of the approximately 3.5 kbp EcoRI/Sall *L. lactis* insert of the plasmid pUCRS deposited under the accession number DSM 5803 and the first HindIII restriction site located in said insert upstream of the base corresponding to position 1 of the DNA sequence with SEQ ID No. 1. up to about
20 the base corresponding to position 491 of the DNA sequence with SEQ ID No. 1.
12. A hybrid vector according to claim 3 in which the functional fragment extends from a base in about position 492 up to about position 1793 of the DNA sequence with SEQ ID No. 1.
- 25 13. A hybrid vector according to claim 1 or 2 comprising a DNA sequence which hybridizes with the approximately 3.5 kbp EcoRI/Sall *L. lactis* insert of the plasmid pUCRS, deposited under the accession number DSM 5803, or with a functional fragment thereof, or such a hybridizing DNA sequence operatively linked to its natural promoter region.
- 30 14. A hybrid vector according to claim 13 comprising a DNA gene sequence hybridizing to the gene encoding the major secretory protein of *L. lactis* or to a functional fragment thereof.
15. A hybrid vector according to claim 1 comprising a DNA sequence which encodes a signal peptide according to claim 1c).
- 35 16. A hybrid vector according to claim 1 or 2 comprising a homologous or heterologous structural gene which is operatively linked in a proper reading frame to the DNA sequence encoding the signal peptide of the major secretory protein and a promoter.
- 40 17. A hybrid vector according to claim 16 encoding desulfatohirudin or the major secretory protein of *L. Lactis* strain LM 0230 as described in SEQ ID No. 1.
18. A DNA molecule which
 - 45 a) is the approximately 3.5 kbp EcoRI/Sall *L. lactis* insert of the plasmid pUCRS deposited under the accession number DSM 5803, or a functional fragment thereof retaining promoter, signal or structural functions, or
 - b) hybridizes with said insert or with a functional fragment thereof, or is such a hybridizing DNA sequence operatively linked to its natural promoter region, or
 - 50 c) is a degenerate sequence of the DNA sequence covered in a) which encodes a signal peptide.
19. A hybrid DNA molecule according to claim 18 comprising the origin of replication of a plasmid belonging to the same incompatibility group as the 2.5 Md or the 5.2 Md plasmid of *L. lactis* strain LL712 deposited under the accession number DSM 5804, which origin can be obtained from the approximately 1.8 kbp long NdeI/SphI fragment of the 2.5 Md plasmid of *L. lactis* LL712 or the 1.0 kbp NdeI/HindIII fragment of the 5.2 Md plasmid of *L. lactis* LL712.
55
20. A host transformed with a hybrid vector according to claim 1 or 2.

21. A transformed host according to claim 20, wherein the host is selected from the group consisting of *E. coli* strains TG1, C600 and HB101; *L. cremoris* strains; plasmid-free *L. lactis* strains and *Bacillus* strains.

22. A transformed host according to claim 20 transformed with a hybrid vector according to claim 17.

23. A method for the preparation of a transformed host according to claim 20 comprising the steps of (a) treating a host under transforming conditions with a hybrid vector of claim 1 or 2, optionally together with a selection marker gene, and (b) selecting the transformants.

24. A method for the preparation of a polypeptide comprising the steps of (a) fusing a homologous or heterologous structural gene in proper reading frame with the DNA sequence coding for the signal peptide according to claim 9, (b) transforming a suitable host with a hybrid vector comprising a fused gene prepared in (a), (c) secreting the polypeptide encoded by said fused gene from the host cells and, when required, (d) isolating the polypeptide from the supernatant according to conventional methods.

25. A method according to claim 24 in which a host selected from the group of hosts consisting of *Lactococcus spec.* and *Bacillus spec.* is used.

26. A method according to claim 24 in which a bacterial host is used which does not secrete in its culture medium a protease which degrades the expressed protein.

27. A method according to claim 24 in which desulfatohirudin is produced.

28. A method according to claim 27 wherein cells of *L. lactis* strain LM0230 as deposited under the accession number DSM 5805 are used.

29. A method according to claim 24 characterized in that (a) the transformed cells are collected from the nutrient medium either in the log phase or in the stationary phase of the culture, (b) they are resuspended in a smaller volume of fresh nutrient medium or suitable buffer solution and (c) they are further incubated before isolating the desired product from the supernatant.

30. A method according to claim 29 characterized in that the transformed cells are resuspended in about 1 % up to 20 % of the original culture volume.

31. A method according to claim 30 wherein the transformed cells are cells from *L. lactis* strain LM0230 as deposited under the accession number DSM 5805 transformed with a hybrid vector according to claim 17 encoding desulfatohirudin and are resuspended in about 1/10 vol. of fresh nutrient medium and are incubated for about 30 min at about 30°C.

32. A method according to claim 24 in which the major secretory protein of *L. lactis* is produced.

33. The major secretory protein of *L. lactis* in pure form having the amino acid given in SEQ ID No. 1.

34. A method for the production of the protein according to claim 33 comprising the steps of (a) precipitating the proteins from the supernatant of *L. lactis* strain LM0230 as deposited under the accession number DSM 5805 with trichloroacetic acid, (b) electrophoresis of the precipitated protein on a SDS polyacrylamide gel, (c) cutting out the region with the major protein band and (d) electroeluting the protein from the gel.

Patentansprüche

1. Hybridvektor, der aufweist:

a) das etwa 3,5-kbp-EcoRI/Sall-*L. lactis*-Insert des Plasmids pUCRS, das unter der Zugangs-Nr. DSM 5803 hinterlegt ist, oder ein funktionelles Fragment davon, das die Promotor-, Signal- oder Strukturfunktionen beibehält, oder

b) eine DNA-Sequenz, die mit dem Insert oder mit einem funktionellen Fragment davon hybridisiert oder eine

derartige hybridisierende DNA-Sequenz, die betriebsbereit an ihre natürliche Promotor-Region gebunden ist, oder

c) eine degenerierte Sequenz der in a) angegebenen Sequenz, die ein Signalpeptid kodiert.

2. Hybridvektor nach Anspruch 1, der den Replikationsursprung eines Plasmids enthält, das zur gleichen Inkompatibilitätsgruppe wie das 2,5-Md- oder das 5,2-Md-Plasmid des *L. lactis*-Stammes LL712, der unter der Zugangs-Nr. DSM 5804 hinterlegt ist, gehört, wobei der Ursprung aus dem etwa 1,8 kbp langen NdeI/SphI-Fragment des 2,5-Md-Plasmids von *L. lactis* LL712 oder dem 1,0-kbp-NdeI/HindIII-Fragment des 5,2-Md-Plasmids von *L. lactis* LL712 erhältlich ist.
3. Hybridvektor nach Anspruch 1 oder 2, der das etwa 3,5-kbp-EcoRI/Sall-*L. lactis*-Insert des Plasmid pUCRS, das unter der Zugangs-Nr. DSM 5803 hinterlegt ist, oder ein funktionelles Fragment davon enthält.
4. Hybridvektor nach Anspruch 1 oder 2, wobei das funktionelle Fragment die Promotorfunktionen des Hauptsekretionsproteingens beibehält.
5. Hybridvektor nach Anspruch 4, wobei das die Promotorfunktion beibehaltende funktionelle Fragment die Region 1500 bp stromaufwärts der das Signalpeptid kodierenden DNA-Sequenz überspannt.
6. Hybridvektor nach Anspruch 5, wobei das die Promotorfunktion beibehaltende funktionelle Fragment eine Region zwischen 100 bp bis zu 1000 bp stromaufwärts der das Signalpeptid kodierenden DNA-Sequenz überspannt.
7. Hybridvektor nach Anspruch 1 oder 2, wobei das funktionelle Fragment die Funktion der Hauptsekretionsprotein-signalsequenz beibehält.
8. Hybridvektor nach Anspruch 7, wobei das die Signalfunktion beibehaltende funktionelle Fragment die in SEQ-ID-Nr. 2 offenbarte 27 Aminosäuren lange Signalsequenz kodiert.
9. Hybridvektor nach Anspruch 7, wobei das die Signalfunktion beibehaltende funktionelle Fragment sich von einer Base bei etwa der Position 411 der DNA-Sequenz mit der SEQ-ID-Nr. 1 bis etwa zur Position 491 erstreckt.
10. Hybridvektor nach Anspruch 3, wobei das funktionelle Fragment die Promotor- und die Signalfunktionen des Hauptsekretionsproteingens beibehält.
11. Hybridvektor nach Anspruch 10, wobei das funktionelle Fragment aus der Gruppe von Fragmenten ausgewählt ist, die sich von einer beliebigen Base zwischen dem EcoRI-Schnittende des etwa 3,5-kbp-EcoRI/Sall-*L. lactis*-Inserts des unter der Zugangs-Nr. DSM 5803 hinterlegten Plasmids pUCRS und der ersten HindIII-Restriktionsstelle, die sich in diesem Insert stromaufwärts von der Base befinden, die Position 1 der DNA-Sequenz mit der SEQ-ID-Nr. 1 entspricht, bis etwa zu der Base, die der Position 491 der DNA-Sequenz mit der SEQ-ID-Nr. 1 entspricht, erstrecken.
12. Hybridvektor nach Anspruch 3, wobei sich das funktionelle Fragment von einer Base bei etwa der Position 492 bis etwa zur Position 1793 der DNA-Sequenz mit SEQ-ID-Nr. 1 erstreckt.
13. Hybridvektor nach Anspruch 1 oder 2, der eine DNA-Sequenz, die mit dem etwa 3,5-kbp-EcoRI/Sall-*L. lactis*-Insert des unter der Zugangs-Nr. DSM 5803 hinterlegten Plasmids pUCRS, oder mit einem funktionellen Fragment davon, hybridisiert oder eine an ihre natürliche Promotor-Region betriebsbereit gebundene derartige hybridisierende DNA-Sequenz enthält.
14. Hybridvektor nach Anspruch 13, der eine DNA-Gensequenz enthält, die mit dem das Hauptsekretionsprotein von *L. lactis* kodierenden Gen oder einem funktionellen Fragment davon hybridisiert.
15. Hybridvektor nach Anspruch 1, der eine DNA-Sequenz enthält, die ein Signalpeptid nach Anspruch 1 c) kodiert.
16. Hybridvektor nach Anspruch 1 oder 2, der ein homologes oder heterologes Strukturgen enthält, das betriebsbereit in einem richtigen Leseraster zu der das Signalpeptid des Hauptsekretionsproteins kodierenden DNA-Sequenz und einem Promotor gebunden ist.

17. Hybridvektor nach Anspruch 16, der Desulfatohirudin oder das Hauptsekretionsprotein des L.-Lactis-Stammes LM 0230, wie in der SEQ-ID-Nr. 1 beschrieben, kodiert.

18. DNA-Molekül, welches:

a) das etwa 3,5-kbp-EcoRI/Sall-L.-lactis-Insert des Plasmids pUCRS, das unter der Zugangs-Nr. DSM 5803 hinterlegt ist, oder ein funktionelles Fragment davon, das die Promotor-, Signal- oder Strukturfunktionen beibehält, ist oder

b) mit dem Insert oder mit einem funktionellen Fragment davon hybridisiert oder eine derartige hybridisierende DNA-Sequenz ist, die betriebsbereit an ihre natürliche Promotor-Region gebunden ist, oder

c) eine degenerierte Sequenz der in a) angegebenen Sequenz, die ein Signalpeptid kodiert, ist.

19. Hybrid-DNA-Molekül nach Anspruch 18, das den Replikationsursprung eines Plasmids enthält, das zur gleichen Inkompatibilitätsgruppe wie das 2,5-Md- oder das 5,2-Md-Plasmid des L.-lactis-Stammes LL712, der unter der Zugangs-Nr. DSM 5804 hinterlegt ist, gehört, wobei der Ursprung aus dem etwa 1,8 kbp langen NdeI/SphI-Fragment des 2,5-Md-Plasmids von L. lactis LL712 oder dem 1,0-kbp-NdeI/HindIII-Fragment des 5,2-Md-Plasmids von L. lactis LL712 erhältlich ist.

20. Wirt, der mit einem Hybridvektor nach Anspruch 1 oder 2 transformiert ist.

21. Transformierter Wirt nach Anspruch 20, wobei der Wirt unter den E.-coli-Stämmen TG1, C600 und HB101, L.-cremoris-Stämmen, plasmidfreien L.-lactis-Stämmen und Bacillus-Stämmen ausgewählt ist.

22. Transformierter Wirt nach Anspruch 20, der mit einem Hybridvektor nach Anspruch 17 transformiert ist.

23. Verfahren zur Herstellung eines transformierten Wirts nach Anspruch 20, bei dem (a) ein Wirt unter transformierenden Bedingungen mit einem Hybridvektor nach Anspruch 1 oder 2, gegebenenfalls zusammen mit einem Selektionsmarkergen, behandelt und (b) die Transformanten selektiert werden.

24. Verfahren zur Herstellung eines Polypeptids, bei dem (a) ein homologes oder heterologes Strukturgen im richtigen Leseraster mit der das Signalpeptid kodierenden DNA-Sequenz nach Anspruch 9 fusioniert, (b) ein geeigneter Wirt mit einem in (a) hergestelltes Fusionsgen enthaltenden Hybridvektor transformiert, (c) das von dem Fusionsgen kodierte Polypeptid von den Wirtszellen sezerniert und, sofern erforderlich, (d) das Polypeptid nach herkömmlichen Verfahren aus dem Überstand isoliert wird.

25. Verfahren nach Anspruch 24, wobei ein unter Lactococcus-Arten und Bacillus-Arten ausgewählter Wirt verwendet wird.

26. Verfahren nach Anspruch 24, wobei ein bakterieller Wirt verwendet wird, der in sein Kulturmedium keine das exprimierte Protein abbauende Protease sezerniert.

27. Verfahren nach Anspruch 24, wobei Desulfatohirudin gebildet wird.

28. Verfahren nach Anspruch 27, wobei Zellen des unter der Zugangs-Nr. DSM 5805 hinterlegten L.-lactis-Stammes LM0230 verwendet werden.

29. Verfahren nach Anspruch 24, dadurch gekennzeichnet, daß

(a) die transformierten Zellen entweder in der log-Phase oder in der stationären Phase der Kultur aus dem Nährmedium gesammelt werden, (b) sie in einem kleineren Volumen frischen Nährmediums oder geeigneter Pufferlösung resuspendiert und (c) sie vor der Isolierung des gewünschten Produktes aus dem Überstand weiter inkubiert werden.

30. Verfahren nach Anspruch 29, dadurch gekennzeichnet, daß

die transformierten Zellen in etwa 1 % bis zu 20 % des ursprünglichen Kulturvolumens resuspendiert werden.

31. Verfahren nach Anspruch 30, wobei die transformierten Zellen Zellen des unter der Zugangs-Nr. DSM 5805 hinterlegten *L.-lactis*-Stammes LM0230 sind, die mit einem Desulfathohirudin kodierenden Hybridvektor nach Anspruch 17 transformiert sind und in etwa 1/10 Vol. frischem Nährmedium resuspendiert und etwa 30 min bei etwa 30°C inkubiert werden.

32. Verfahren nach Anspruch 24, wobei das Hauptsekretionsprotein von *L. lactis* gebildet wird.

33. Hauptsekretionsprotein von *L. lactis* in reiner Form mit der in der SEQ-ID-Nr. 1 angegebenen Aminosäuresequenz.

34. Verfahren zur Herstellung des Proteins nach Anspruch 33, bei dem (a) die Proteine aus dem Überstand des unter der Zugangs-Nr. DSM 5805 hinterlegten *L.-lactis*-Stammes LM0230 mit Trichloressigsäure ausgefällt, (b) das ausgefällte Protein auf einem SDS-Polyacrylamidgel elektrophoretisiert, (c) die Region mit der Hauptproteinbande ausgeschnitten und (d) das Protein aus dem Gel elektroeluiert wird.

Revendications

1. Vecteur hybride comprenant

- a) le segment d'insertion de *L. lactis* EcoRI/Sall d'approximativement 3,5 kpb du plasmide pUCRS, déposé sous le n° de dépôt DSM 5803, ou un de ses fragments fonctionnels conservant les fonctions de promoteur, de signal ou des fonctions structurales, ou
- b) une séquence d'ADN qui s'hybride avec ledit segment d'insertion ou avec un fragment fonctionnel de ce segment, ou une telle séquence d'ADN hybridante liée de manière fonctionnelle à sa région de promoteur naturelle, ou
- c) une séquence dégénérée d'une séquence d'ADN qui répond à la définition en a) et qui code pour un peptide-signal.

2. Vecteur hybride suivant la revendication 1, comprenant l'origine de répllication d'un plasmide appartenant au même groupe d'incompatibilité que le plasmide de 2,5 Md ou de 5,2 Md de la souche *L. lactis* LL712 déposée sous le n° de dépôt DSM 5804, origine qui peut être obtenue à partir du fragment NdeI/SphI d'approximativement 1,8 kpb de longueur du plasmide de 2,5 Md de *L. lactis* LL712 ou du fragment NdeI/HindIII de 1,0 kpb du plasmide de 5,2 Md de *L. lactis* LL712.

3. Vecteur hybride suivant la revendication 1 ou 2, comprenant le segment d'insertion de *L. lactis* EcoRI/Sall d'approximativement 3,5 kpb du plasmide pUCRS déposé sous le n° de dépôt DSM 5803, ou un de ses fragments fonctionnels.

4. Vecteur hybride suivant la revendication 1 ou 2, dans lequel le fragment fonctionnel conserve les fonctions de promoteur du gène de protéine sécrétoire majeure.

5. Vecteur hybride suivant la revendication 4, dans lequel le fragment fonctionnel conservant la fonction de promoteur s'étend sur la région de 1500 pb en amont de la séquence d'ADN codant pour le peptide-signal.

6. Vecteur hybride suivant la revendication 5, dans lequel le fragment fonctionnel conservant la fonction de promoteur s'étend sur une région allant de 100 pb jusqu'à 1000 pb en amont de la séquence d'ADN codant pour le peptide-signal.

7. Vecteur hybride suivant la revendication 1 ou 2, dans lequel le fragment fonctionnel conserve la fonction de la séquence signal de la protéine sécrétoire majeure.

8. Vecteur hybride suivant la revendication 7, dans lequel le fragment fonctionnel conservant la fonction de signal code pour la séquence-signal de 27 amino-acides de longueur décrite dans la SEQ ID n° 2.

9. Vecteur hybride suivant la revendication 7, dans lequel le fragment fonctionnel conservant la fonction de signal s'étend d'une base approximativement à la position 411 de la séquence d'ADN ayant la SEQ ID n° 1 jusqu'à approximativement la position 491.

10. Vecteur hybride suivant la revendication 3, dans lequel le fragment fonctionnel conserve la fonction de promoteur et la fonction de signal du gène de la protéine sécrétoire majeure.
- 5 11. Vecteur hybride suivant la revendication 10, dans lequel le fragment fonctionnel est choisi dans le groupe de fragments s'étendant de n'importe quelle base entre l'extrémité coupée EcoRI du segment d'insertion de *L. lactis* EcoRI/Sall d'approximativement 3,5 kpb du plasmide pUCRS déposé sous le n° de dépôt DSM 5803 et le premier site de restriction HindIII situé dans ledit segment d'insertion en amont de la base correspondant à la position 1 de la séquence d'ADN ayant la SEQ ID n° 1 jusqu'à approximativement la base correspondant à la position 491 de la séquence d'ADN ayant la SEQ ID n° 1.
- 10 12. Vecteur hybride suivant la revendication 3, dans lequel le fragment fonctionnel s'étend d'une base approximativement à la position 492 jusqu'à approximativement la position 1793 de la séquence d'ADN ayant la SEQ ID n° 1.
- 15 13. Vecteur hybride suivant la revendication ou 2, comprenant une séquence d'ADN qui s'hybride avec le segment d'insertion de *L. lactis* EcoRI/Sall d'approximativement 3,5 kpb du plasmide pUCRS, déposé sous le n° de dépôt DSM 5803, ou avec un de ses fragments fonctionnels, ou bien une telle séquence d'ADN hybridante liée de manière fonctionnelle à sa région de promoteur naturelle.
- 20 14. Vecteur hybride suivant la revendication 13, comprenant une séquence de gène d'ADN s'hybridant au gène codant pour la protéine sécrétoire majeure de *L. lactis* ou à un de ses fragments fonctionnels.
- 25 15. Vecteur hybride suivant la revendication 1, comprenant une séquence d'ADN qui code pour un peptide-signal suivant la revendication 1c).
- 30 16. Vecteur hybride suivant la revendication 1 ou 2, comprenant un gène structural homologue ou hétérologue, qui est lié de manière fonctionnelle en cadre de lecture convenable à la séquence d'ADN codant pour le peptide-signal de la protéine sécrétoire majeure et un promoteur.
- 35 17. Vecteur hybride suivant la revendication 16, codant pour la désulfatohirudine ou la protéine sécrétoire majeure de la souche *L. lactis* LM0230 décrite dans la SEQ ID n° 1.
- 40 18. Molécule d'ADN qui
 - a) est le segment d'insertion de *L. lactis* EcoRI/Sall d'approximativement 3,5 kpb du plasmide pUCRS, déposé sous le n° de dépôt DSM 5803, ou un de ses fragments fonctionnels conservant les fonctions de promoteur, de signal ou des fonctions structurales, ou
 - b) s'hybride avec ledit segment d'insertion ou avec un fragment fonctionnel de ce segment, ou bien est une telle séquence d'ADN hybridante liée de manière fonctionnelle à sa région de promoteur naturelle, ou
 - c) est une séquence dégénérée d'une séquence d'ADN qui est incluse dans a) et qui code pour un peptide-signal.
- 45 19. Molécule d'ADN hybride suivant la revendication 18, comprenant l'origine de répllication d'un plasmide appartenant au même groupe d'incompatibilité que le plasmide de 2,5 Md ou 5,2 Md de la souche *L. lactis* LL712 déposée sous le n° de dépôt DSM 5804, origine qui peut être obtenue à partir du fragment NdeI/SphI d'approximativement 1,8 kpb de longueur du plasmide de 2,5 Md de *L. lactis* LL712 ou du fragment NdeI/HindIII d'approximativement 1,0 kpb de longueur du plasmide de 5,2 Md de *L. lactis* LL712.
- 50 20. Hôte transformé avec un vecteur hybride suivant la revendication 1 ou 2.
- 55 21. Hôte transformé suivant la revendication 20, dans lequel l'hôte est choisi dans le groupe consistant en les souches *E. coli* TG1, C600 et HB101 ; des souches de *L. cremoris* ; des souches de *L. lactis* et des souches de *Bacillus* dépourvues de plasmides.
22. Hôte transformé suivant la revendication 20, qui a été transformé avec un vecteur hybride suivant la revendication 17.
23. Procédé pour la préparation d'un hôte transformé suivant la revendication 20, comprenant les étapes consistant (a) à traiter un hôte dans des conditions de transformation avec un vecteur hybride suivant la revendication 1 ou

2, facultativement en association avec un gène marqueur de sélection, et (b) à sélectionner les transformants.

24. Procédé pour la transformation d'un polypeptide, comprenant les étapes consistant (a) à fusionner un gène structural homologue ou hétérologue en cadre de lecture convenable avec la séquence d'ADN codant pour le peptide-signal suivant la revendication 9, (b) à transformer un hôte convenable avec un vecteur hybride comprenant un gène fusionné préparé en (a), (c) à provoquer la sécrétion du polypeptide codé par ledit gène de fusion par les cellules hôtes et, lorsque cela est requis, (d) à isoler le polypeptide du surnageant par des procédés classiques.

25. Procédé suivant la revendication 24, dans lequel un hôte choisi dans le groupe d'hôtes consistant en *Lactococcus spec.* et *Bacillus spec.* est utilisé.

26. Procédé suivant la revendication 24, dans lequel est utilisé un hôte bactérien qui ne sécrète pas dans son milieu de culture une protéase qui dégrade la protéine exprimée.

27. Procédé suivant la revendication 24, dans lequel de la désulfatohirudine est produite.

28. Procédé suivant la revendication 27, dans lequel des cellules de la souche *L. lactis* LM0230, déposée sous le n° de dépôt DSM 5805, sont utilisées.

29. Procédé suivant la revendication 24, caractérisé en ce que (a) les cellules transformées sont collectées à partir du milieu nutritif en phase logarithmique ou en phase stationnaire de la culture, (b) elles sont remises en suspension dans un plus petit volume de milieu nutritif frais ou de solution tampon convenable, et (c) elles sont en outre mises en incubation avant d'isoler le produit désiré du surnageant.

30. Procédé suivant la revendication 29, caractérisé en ce que les cellules transformées sont remises en suspension dans un volume d'environ 1 % jusqu'à 20 % du volume de culture initial.

31. Procédé suivant la revendication 30, dans lequel les cellules transformées sont des cellules de la souche *L. lactis* LM0230, déposée sous le n° de dépôt DSM 5805 transformées avec un vecteur hybride suivant la revendication 17 codant pour la désulfatohirudine et sont remises en suspension dans approximativement 1/10 de volume de milieu nutritif frais et sont mises en incubation pendant un temps d'environ 30 minutes à une température d'environ 30 °C.

32. Procédé suivant la revendication 24, dans lequel la protéine sécrétoire majeure de *L. lactis* est produite.

33. Protéine sécrétoire majeure de *L. lactis* sous forme pure, ayant la séquence d'acides-amino mentionnée dans la SEQ ID n° 1.

34. Procédé pour la production de la protéine suivant la revendication 33, comprenant les étapes consistant (a) à précipiter les protéines à partir du surnageant de la souche *L. lactis* LM0230, déposée sous le n° de dépôt DSM 5805, avec de l'acide trichloracétique, (b) à soumettre à une électrophorèse la protéine précipitée sur un gel de polyacrylamide au SDS, (c) à exciser la région portant la bande de protéine principale, et (d) à électroéluer la protéine hors du gel.

Figure 1

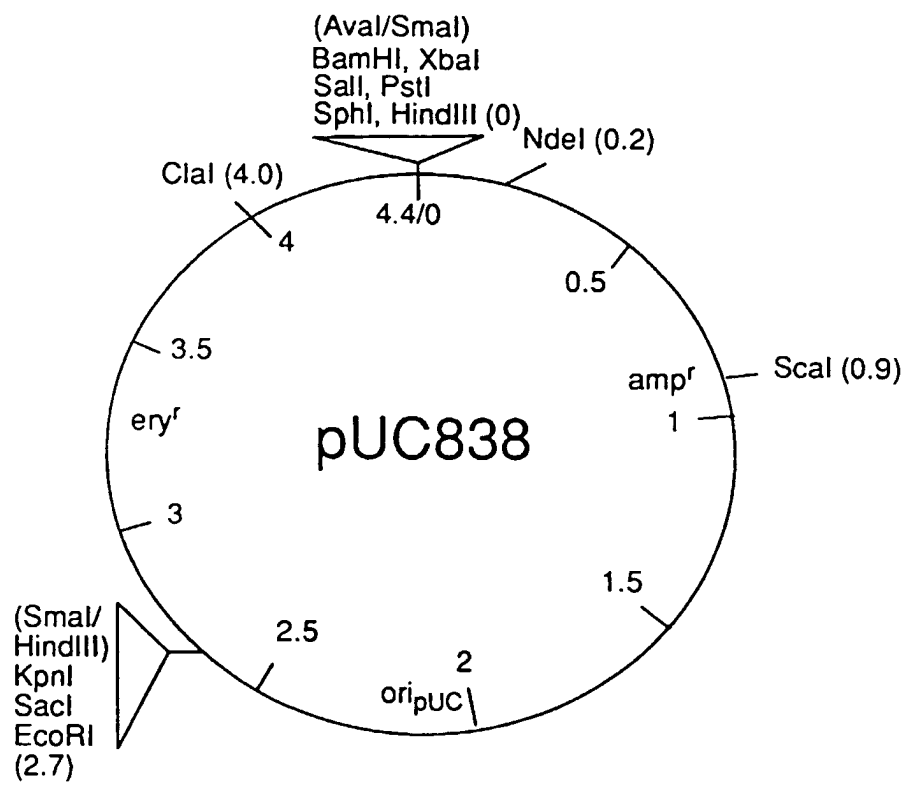


Figure 2

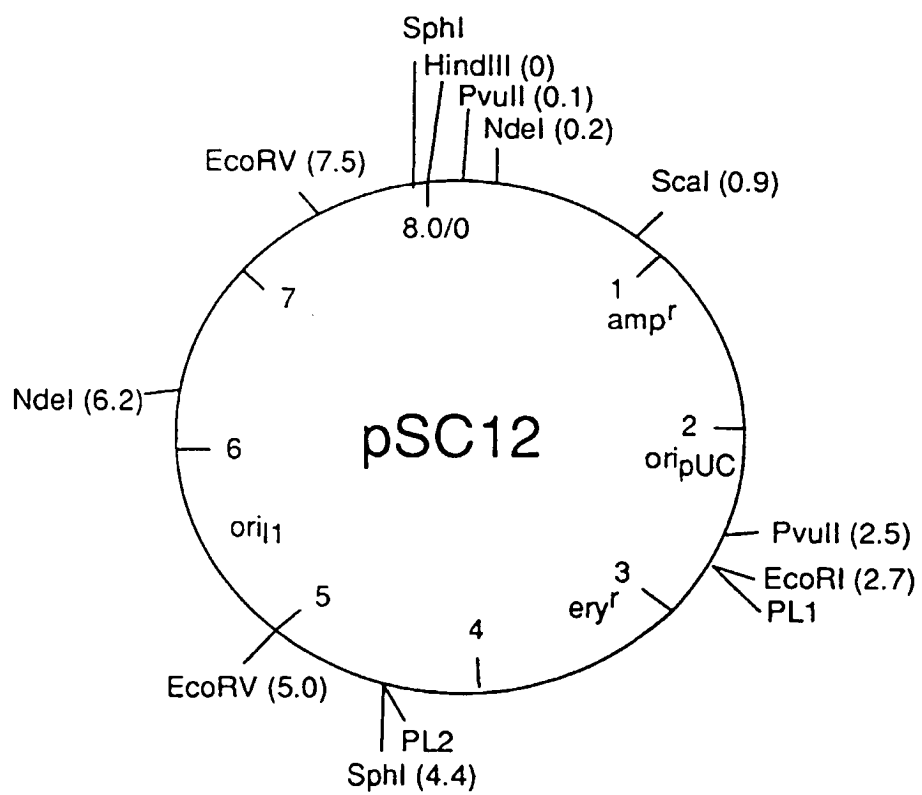


Figure 3

